

Universidad Politécnica de Cartagena

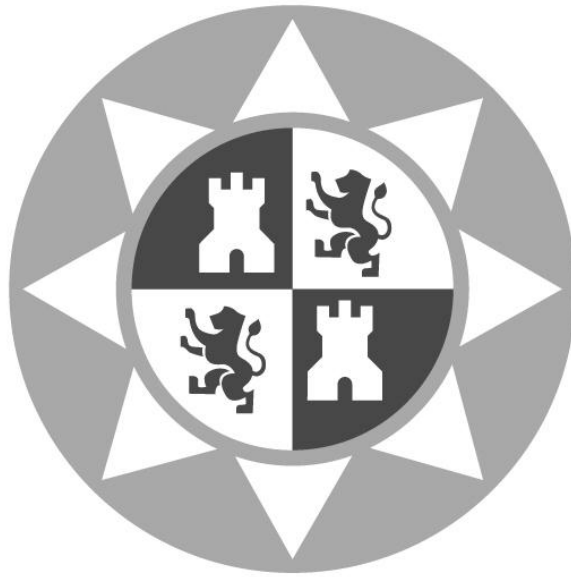
Departamento de Tecnologías de la Información y las Comunicaciones

Contribution to Quality of user Experience Provision over Wireless Networks

Doctoral Thesis

Ramón J. Sánchez Iborra

2015



Universidad Politécnica de Cartagena

Departamento de Tecnologías de la Información y las Comunicaciones

Contribution to Quality of user Experience provision over Wireless Networks

Ramón J. Sánchez Iborra

Advisor:

María Dolores Cano Baños

2015



**CONFORMIDAD DE SOLICITUD DE AUTORIZACIÓN DE DEPÓSITO DE TESIS
DOCTORAL POR EL/LA DIRECTOR/A DE LA TESIS**

D./D^a. María Dolores Cano Baños, Director/a de la Tesis doctoral "Contribution to Quality of User Experience Provision over Wireless Networks"

INFORMA:

Que la referida Tesis Doctoral, ha sido realizada por D/D^a. Ramón Jesús Sánchez Iborra, dentro del programa de doctorado "Tecnologías de la Información y las Comunicaciones, dando mi conformidad para que sea presentada ante la Comisión de Doctorado para ser autorizado su depósito.

La rama de conocimiento en la que esta tesis ha sido desarrollada es:

- ☐ Ciencias
- ☐ Ciencias Sociales y Jurídicas
- ☒ Ingeniería y Arquitectura

En Cartagena, a 4 de Septiembre de 2015

EL/LA DIRECTOR/A DE LA TESIS

Fdo. _____

COMISIÓN DE DOCTORADO



**CONFORMIDAD DE DEPÓSITO DE TESIS DOCTORAL
POR LA COMISIÓN ACADÉMICA DEL PROGRAMA**

D/D^a. Fernando Daniel Quesada Pereira, Presidente/a de la Comisión Académica del Programa de Doctorado “Tecnologías de la Información y las Comunicaciones”.

INFORMA:

Que la Tesis Doctoral titulada, “Contribution to Quality of User Experience Provision over Wireless Networks”, ha sido realizada, dentro del mencionado programa de doctorado, por D/D^a. Ramón Jesús Sánchez Iborra, bajo la dirección y supervisión del Dr/ Dra. María Dolores Cano Baños.

En reunión de la Comisión Académica de fecha **7 /09/15**, visto que en la misma se acreditan los indicios de calidad correspondientes y la autorización del Director de la misma, se acordó dar la conformidad, con la finalidad de que sea autorizado su depósito por la Comisión de Doctorado.

La Rama de conocimiento por la que esta tesis ha sido desarrollada es:

Ciencias
Ciencias Sociales y Jurídicas
☒ Ingeniería y Arquitectura

En Cartagena, a **7** de Septiembre de 2015

EL PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA

Fdo. 



COMISIÓN DE DOCTORADO

Agradecimientos

Una Tesis doctoral, como toda etapa en la vida, conlleva un aprendizaje a muchos niveles. Esta Tesis, cuyo desarrollo no siempre ha sido fácil, me ha permitido evolucionar y crecer, tanto en el plano profesional como personal. Por suerte, este camino no lo he recorrido solo y son muchas personas a las que debo agradecer todo lo que he aprendido de ellas.

En primer lugar, debo agradecer a María Dolores Cano (Lola) el haberme dado la oportunidad de iniciarme en el apasionante mundo investigador. Son muchos años trabajando juntos y estoy convencido de que no hubiera llegado hasta aquí siendo dirigido por otra persona. Gracias por tu optimismo y tus inagotables ideas felices. Espero que el final de esta Tesis solo suponga un punto y seguido y sigamos colaborando por mucho tiempo.

A Joan García Haro, por la confianza depositada en mí y su apoyo incondicional a todos los niveles. Gracias por tus concienzudas revisiones, comentarios realistas y, sobre todo, por enseñarme como debo comportarme si algún día tengo gente a mi cargo.

Al grupo de investigación NetGNA de la Universidad de Beira Interior y, en especial, al Dr. Joel Rodrigues por recibirme de la mejor manera y hacerme sentir como un miembro más del equipo desde el primer día. Muchas gracias Joel por mostrarme en el ya lejano ICC'13 “el otro mundo de la investigación”. A mis compañeros “portunholes” de laboratorio: Sandra, Edgar, Susana, Pedro Pinto y Gilberto, por vuestro apoyo durante mis interminables sesiones de “ventanas negras” y, sobre todo, por los buenos momentos fuera del laboratorio.

También me gustaría agradecer al Quality and Usability Lab de la TU Berlín, encabezado por el Dr. Sebastian Möller, por aceptar que formara parte de uno de los grupos de investigación *top* en el mundo. Gracias por enseñarme otras formas de trabajar y por estar siempre dispuesto a colaborar. A mis compañeros Falk, Jan-Niklas y Sebastian por permitirme participar en sus apasionantes proyectos y poder aportar mi granito de arena en ellos. Mención especial para Laura, por acogerme desde antes de que llegara e integrarme en su grupo de españoles expatriados.

Al Dr. Alfonso Ariza de la Universidad de Málaga, el cual probablemente no leerá jamás estas líneas, pues no me conoce de nada y aun así, ha respondido de forma incansable a mis

cuestiones sobre el simulador. Gracias por esta muestra de colaboración altruista y completamente desinteresada.

A Marcos Abad, incombustible técnico de laboratorio al que le caen “todos los marrones” y que, pese a ello, ha estado siempre dispuesto a echar un mano en lo que fuera necesario. También por esas conversaciones en las que arreglábamos el mundo desde Jumilla.

A mis amigos, los POT y el grupo de Mazarrón, por esos ratos de desconexión, en los que se habla de cualquier cosa menos de la Tesis. Hemos crecido juntos y lo seguiremos haciendo allá donde estemos.

A mi familia, los que están, abuela, tíos y primos, y los que nos dejaron, por su interés y cariño constantes hacia mí.

A Marta, gracias por compartir tu vida conmigo y por hacer que todo resultara un poco más fácil. Gracias por tus palabras cuando eran necesarias y por ayudarme a ver el mundo con perspectiva cuando parecía que las cosas no salían.

A mis padres, por sus sabios consejos y por estar siempre ahí. Sin vuestro apoyo no sería quien soy ni hubiera llegado hasta aquí. Aunque sigáis empeñados en que “invente algo y me haga rico”, estoy seguro de que ese final no se alcanza sin haber recorrido antes un largo camino. Gracias por los valores que me habéis transmitido y por ser un ejemplo de constancia, trabajo y esfuerzo.

A mis padres, Ramón y Bárbara

*“Any sufficiently advanced technology is
indistinguishable from magic”*

Arthur C. Clarke

Abstract

The widespread expansion of wireless networks has brought new attractive possibilities to end users. In addition to the mobility capabilities provided by unwired devices, it is worth remarking the easy configuration process that a user has to follow to gain connectivity through a wireless network. Furthermore, the increasing bandwidth provided by the IEEE 802.11 family has made possible accessing to high-demanding services such as multimedia communications. Multimedia traffic has unique characteristics that make it greatly vulnerable against network impairments, such as packet losses, delay, or jitter. Voice over IP (VoIP) communications, video-conference, video-streaming, etc., are examples of these high-demanding services that need to meet very strict requirements in order to be served with acceptable levels of quality. Accomplishing these tough requirements will become extremely important during the next years, taking into account that consumer video traffic will be the predominant traffic in the Internet during the next years.

In wired systems, these requirements are achieved by using Quality of Service (QoS) techniques, such as Differentiated Services (DiffServ), traffic engineering, etc. However, employing these methodologies in wireless networks is not that simple as many other factors impact on the quality of the provided service, e.g., fading, interferences, etc. Focusing on the IEEE 802.11g standard, which is the most extended technology for Wireless Local Area Networks (WLANs), it defines two different architecture schemes. On one hand, the infrastructure mode consists of a central point, which manages the network, assuming network controlling tasks such as IP assignment, routing, accessing security, etc. The rest of the nodes composing the network act as hosts, i.e., they send and receive traffic through the central point. On the other hand, the IEEE 802.11 ad-hoc configuration mode is less extended than the infrastructure one. Under this scheme, there is not a central point in the network, but all the nodes composing the network assume both host and router roles, which permits the quick deployment of a network without a pre-existent infrastructure. This type of networks, so called Mobile Ad-hoc NETWORKs (MANETs), presents interesting characteristics for situations when the fast deployment of a communication system is needed, e.g., tactics networks, disaster events, or temporary networks. The benefits provided by MANETs are varied, including high mobility possibilities provided to the nodes, network

coverage extension, or network reliability avoiding single points of failure. The dynamic nature of these networks makes the nodes to react to topology changes as fast as possible. Moreover, as aforementioned, the transmission of multimedia traffic entails real-time constraints, necessary to provide these services with acceptable levels of quality. For those reasons, efficient routing protocols are needed, capable of providing enough reliability to the network and with the minimum impact to the quality of the service flowing through the nodes.

Regarding quality measurements, the current trend is estimating what the end user actually perceives when consuming the service. This paradigm is called Quality of user Experience (QoE) and differs from the traditional Quality of Service (QoS) approach in the human perspective given to quality estimations. In order to measure the subjective opinion that a user has about a given service, different approaches can be taken. The most accurate methodology is performing subjective tests in which a panel of human testers rates the quality of the service under evaluation. This approach returns a quality score, so-called Mean Opinion Score (MOS), for the considered service in a scale 1 - 5. This methodology presents several drawbacks such as its high expenses and the impossibility of performing tests at real time. For those reasons, several mathematical models have been presented in order to provide an estimation of the QoE (MOS) reached by different multimedia services

In this thesis, the focus is on evaluating and understanding the multimedia-content transmission-process in wireless networks from a QoE perspective. To this end, firstly, the QoE paradigm is explored aiming at understanding how to evaluate the quality of a given multimedia service. Then, the influence of the impairments introduced by the wireless transmission channel on the multimedia communications is analyzed. Besides, the functioning of different WLAN schemes in order to test their suitability to support highly demanding traffic such as the multimedia transmission is evaluated.

Finally, as the main contribution of this thesis, new mechanisms or strategies to improve the quality of multimedia services distributed over IEEE 802.11 networks are presented. Concretely, the distribution of multimedia services over ad-hoc networks is deeply studied. Thus, a novel opportunistic routing protocol, so-called JOKER (auto-adJustable Opportunistic acK/timEr-based Routing) is presented. This proposal permits better support to multimedia services while reducing the energy consumption in comparison with the standard ad-hoc routing protocols.

Contents

ACRONYMS AND ABBREVIATIONS	VII
CHAPTER 1 Introduction	1
1.1 Presentation	1
1.2 Thesis Objectives	3
1.3 Dissertation Organization	3
1.4 Thesis Contributions	5
CHAPTER 2 Quality of Experience in Multimedia Services: Evaluation Techniques	9
2.1 Introduction	9
2.2 Quality of Service and Quality of Experience	10
2.3 VoIP QoE Estimation Methodologies	14
2.3.1 Standard Models	16
2.3.1.1 E-Model	16
2.3.1.2 Improvements to the E-Model	19
2.3.1.3 Call Clarity Index	25
2.3.1.4 P.563	27
2.3.1.5 ANIQUE+	30
2.3.1.6 P.564	31
2.3.2 Non-Standard Models	32
2.3.2.1 Gaussian Mixture Models	33
2.3.2.2 Neural Networks	34
2.3.2.3 Exponential Functions	36
2.3.2.4 Genetic Programming	37
2.3.2.5 Hybrids Models	37
2.3.2.6 Other Methodologies	39
2.4 IPTV QoE Estimation Methodologies	42
2.4.1 Standard Models	43
2.4.2 Non-Standard Models	53
2.4.2.1 Audiovisual-only Quality Assessment Models	54
2.4.2.2 Complete-user-experience Quality Assessment Models	64
2.5 Case Study: Evaluating the Quality of VoIP Services in a Real Environment	67
2.5.1 Test-Bench and Components Features	67
2.5.1.1 Locutions	67
2.5.1.2 Transmitter and Receiver Equipment	68
2.5.1.3 Participants	69
2.5.1.4 Experimental Methodology	69
2.5.2 QoE Data Collection	70
2.5.2.1 Questionnaire	70
2.5.2.2 Calculations	70

2.5.3 QoS Data Collection	71
2.5.4 QoE Results	72
2.5.5 QoS Results	79
2.5.5.1 Bandwidth	79
2.5.5.2 Packet Size	80
2.5.5.3 Jitter	81
2.5.5.4 Packet Loss	82
2.5.5.5 Non-VoIP Bandwidth	82
2.5.5.6 Delay	83
2.6 Conclusion	84
2.7 Contributions Related to this Chapter	86
CHAPTER 3 Performance Evaluation of 802.11g Infrastructure Networks Supporting VoIP Traffic: Impact of the PHY Layer and Voice Codec on QoE	87
3.1 Introduction	87
3.2 Propagation Models	88
3.2.1 Free Space Model	89
3.2.2 Rayleigh Model	91
3.2.3 Rice Model	93
3.2.4 Nakagami- <i>m</i> Model	94
3.3 VoIP QoE Evaluation in 802.11g Infrastructure Networks	95
3.3.1 System Theoretical Capacity	97
3.3.2 E-Model Implementation	98
3.3.3 Simulation Environment	100
3.3.4 Results	103
3.3.4.1 Regular Network Evaluation	103
3.3.4.1.1 Packetization of 20 ms	104
3.3.4.1.2 Packetization of 10 ms	111
3.3.4.2 Fading Impact Evaluation	112
3.3.4.2.1 Packetization of 10 ms	112
3.3.4.2.2 Packetization of 20 ms	120
3.4 Conclusion	120
3.5 Contributions Related to this Chapter	122
CHAPTER 4 Performance Evaluation of 802.11g Mobile Ad-Hoc Networks Supporting Multimedia Traffic: a QoE Perspective	123
4.1 Introduction	123
4.2 Routing Protocols in MANETs	125
4.2.1 OLSR	126
4.2.2 BATMAN	127
4.3 Test-Benches	128
4.4 Results	132

4.4.1 VoIP Traffic	132
4.4.1.1 Impact of Fading	133
4.4.1.2 Impact of Number and Density of Nodes	138
4.4.1.3 Impact of Nodes Mobility	143
4.4.2 Video-streaming Traffic	147
4.4.2.1 Regular Performance Evaluation	148
4.4.2.2 Tuning of Routing Messages Intervals: Impact on the VoIP Service	153
4.4.2.3 Establishing the Maximum Video Bit-rate	161
4.4.2.4 Tuning of Routing Messages Intervals: Impact on the Video-streaming Service	162
4.5 Conclusion	167
4.6 Contributions Related to this Chapter	169
CHAPTER 5 Novel Proposals for Improving QoE and Energy Efficiency in 802.11g Mobile Ad-Hoc Networks	171
5.1 Introduction	171
5.2 Energy Consumption vs. QoE	174
5.2.1 Recent Advances in Energy-efficiency in Ad-hoc Networks	174
5.3 Opportunistic Routing	176
5.3.1 Recent Advances in Opportunistic Routing	178
5.4 JOKER	181
5.4.1 Main Characteristics	182
5.4.2 Candidate Selection	183
5.4.3 Candidate Coordination	186
5.5 Development Details & Test-Bench Description	187
5.5.1 Simulation Test-bed	187
5.5.2 Experimental Test-bed	189
5.6 Results	191
5.6.1 QoE - Energy Consumption Trade-off Evaluation	192
5.6.1.1 QoE vs. Energy Consumption Study of BATMAN, OLSR, and AODV	192
5.6.1.2 Improving BATMAN performance with dynamic OGM intervals: Dyn-BATMAN	196
5.6.2 JOKER Evaluation	202
5.6.2.1 Protocol Adjustment	202
5.6.2.2 JOKER vs. BATMAN	210
5.7 Conclusion	217
5.8 Contributions Related to this Chapter	219
CHAPTER 6 Conclusion	221
6.1 Conclusion	221
REFERENCES	225

Acronyms and Abbreviations

ACK	<i>Acknowledgment</i>
ACR	<i>Absolute Category Rating</i>
ADPCM	<i>Adaptive Differential Pulse Code Modulation</i>
ANOVA	<i>Analysis of Variance</i>
AODV	<i>Ad hoc On-Demand Distance Vector</i>
BATMAN	<i>Better Approach To Mobile Adhoc Networking</i>
BER	<i>Bit Error Rate</i>
CBR	<i>Constant Bit Rate</i>
CIF	<i>Common Intermediate Format</i>
CMSI	<i>Control Message Sending Interval</i>
CTS	<i>Clear To Send</i>
DiffServ	<i>Differentiated Services</i>
ETX	<i>Expected Transmissions Count</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IETF	<i>Internet Engineering Task Force</i>
IP	<i>Internet Protocol</i>
IPTV	<i>Internet Protocol Television</i>
ITU-T	<i>International Telegraph Union - Telecommunication Standardization Sector</i>
JOKER	<i>auto-adJustable Opportunistic ACK/timEr-based Routing</i>
Kbps	<i>Kilobits per second</i>
LQ	<i>Link Quality</i>
m	<i>meter</i>
MAC	<i>Media Access Control</i>
MANET	<i>Mobile Ad-hoc Network</i>

MOS	<i>Mean Opinion Score</i>
MOS-CQ	<i>MOS Conversational Quality</i>
MOS-LQ	<i>MOS Listening Quality</i>
MPR	<i>Multi Point Relaying</i>
ms	<i>milisecond</i>
OGM	<i>Originator Message</i>
OLSR	<i>Optimized Link State Routing</i>
PCM	<i>Pulse Code Modulation</i>
P-E-model	<i>Packet E-model</i>
PESQ	<i>Perceptual Evaluation of Speech Quality</i>
PLR	<i>Packet Loss Ratio</i>
Ppl	<i>Probability of Packet Loss</i>
QCIF	<i>Quarter Common Intermediate Format</i>
QoE	<i>Quality of Experience</i>
QoS	<i>Quality of Service</i>
RFC	<i>Request for Comments</i>
RTP	<i>Real Time Protocol</i>
RTS	<i>Ready To Send</i>
RTT	<i>Round Trip Time</i>
RX	<i>Receiver</i>
s	<i>second</i>
SLA	<i>Service Level Agreement</i>
SNIR	<i>Signal to Interference and Noise Ratio</i>
TCP	<i>Transmission Control Protocol</i>
TQ	<i>Transmission Quality</i>
TX	<i>Transmitter</i>
UDP	<i>User Datagram Protocol</i>

VBR	<i>Variable Bit Rate</i>
VoIP	<i>Voice over IP</i>
VoIP-EM	<i>VoIP E-model</i>
Wi-Fi	<i>Wireless Fidelity</i>
WLAN	<i>Wireless Local Area Network</i>
WSN	<i>Wireless Sensor Network</i>

Chapter 1

Introduction

1.1 Presentation

Multimedia transmissions over packet-switched networks are becoming more and more popular among end-users [1]. These types of communications are opening a new wave of unimaginable services little time ago. However, multimedia traffic poses strict requirements to meet in order to provide these services with acceptable levels of quality due to their intensive use of bandwidth. Delay-sensitive services such as Voice over IP (VoIP) or video-streaming, which are some of the most prevalent applications, are also strongly impacted by other impairments such as packet loss or signal distortions. Therefore, mechanisms to ensure the appropriate quality of these services are indispensable. The classical approach to evaluate the quality of a given service, i.e., Quality of Service (QoS), is being replaced by a user-centric approach known as Quality of user Experience (QoE). Unlike the former, which focuses on monitoring network metrics, such as delay, jitter, or Packet Loss Ratio (PLR), the latter quantifies what the user actually perceives when consuming the service. The regular subjective QoE evaluation is performed throughout a quality survey passed to a group of observers, which rate the service in a scale ranging from 1 (poor quality) to 5 (excellent quality) (please, see ITU-T Rec. P.800 [2]). This quality metric, so-called Mean Opinion Score (MOS), is extensively used by the research community to assess the quality level of multimedia services. However, quality surveys are expensive, time-consuming, and cannot be carried out in real time; for these reasons, some mathematical models have been derived aiming to mimic/characterize the quality perception of a human being by means of analyzing technical parameters extracted from the multimedia-flow or the delivery system. These techniques are usually referred to as objective methodologies [3], [4].

Focusing on the customer side, it is unquestionable that wireless networks have brought a new era of experiences and facilities to the end-users. Mobility or ease-of-access, among others, are good examples of greatly valued characteristics by users. Nevertheless, achieving the desired service's level of quality in wireless networks is not

as simple as it is in wired environments. In the latter, most of the system variables are under control and many correcting actions can be taken when the quality of the service drops under certain low values. On the contrary, wireless systems have inherent characteristics related to the transmission medium (physical layer) that make them unpredictable and difficult to maintain under perfect control. New impairments that do not exist in wired networks, e.g., interferences, fading, channel access contention, etc. together with the mobility of the nodes composing the network, pose new difficulties to the successful establishment and support of demanding transmissions, such as multimedia communications.

Currently, the most widespread wireless standard to configure Wireless Local Area Networks (WLANs) is the IEEE 802.11 family (Wi-Fi). This technology defines two different operational modes. The most extended is the infrastructure mode that consists of a central node (i.e., access point) that provides network services to the rest of the network nodes, e.g., routing, IP configuration, etc. The remaining nodes act as hosts sending and receiving data traffic through the central access point. The other IEEE 802.11 configuration scheme, referred to as ad-hoc mode, offers a completely decentralized architecture, extremely attractive for mobility purposes, and very suitable under situations where it is not possible to deploy a centralized transmission system, such as in the event of disasters or the need of setting up temporary networks. In other words, Mobile Ad-hoc NETWORKs (MANETs) can be deployed at any place in a very short time, providing access to information and services regardless of geographic position [5], [6]. Thus, MANETs are considered as a future networking facility, which will allow replacing the expensive infrastructure [7], [8]. Due to the dynamic nature of MANETs, they are able to form a self-reconfigurable network in which all their nodes collaborate in the routing tasks; consequently, nodes in the network can perform the roles of both hosts and routers. By using appropriate routing protocols, all these tasks are transparent to the end-user, which facilitates the deployment of these systems. However, MANETs pose additional challenges, especially related to node mobility and energy constraints. On the one hand, the dynamic multi-hop topology of the network may cause constant route changes, hidden/exposed terminals problems, etc. On the other hand, as the terminals employed by the end-users are battery-powered, the routing tasks should be as computational-friendly as possible. Thus, efficient routing protocols for MANETs are necessary, capable of providing enough reliability to the network

without excessive power consumption, and with the minimum impact on the service data flowing through the nodes.

1.2 Thesis Objectives

The main objective of this thesis is evaluating and understanding the multimedia-content transmission-process in wireless networks aiming at presenting real proposals to improve the quality of these services. This issue presents many dimensions to study: (i) how to evaluate the quality of a given multimedia service, (ii) analyzing the influence of the impairments introduced by the wireless transmission channel on the multimedia communications, or (iii) researching about the functioning of different WLAN schemes in order to test their suitability to support highly demanding traffic such as the multimedia transmission.

To this end, the workflow of this thesis was defined to follow the next stages:

1. Defining concepts such as Quality of Service (QoS) or Quality of user Experience (QoE) and understanding the current approaches employed to assess the quality of multimedia services.
2. Compiling the different methodologies presented in the literature to evaluate the quality of multimedia communications (VoIP and Future TV services) from the end-user perspective.
3. Performing real experiments to analyze the correlation between different-nature quality estimation perspectives (subjective tests vs. objective metrics).
4. Evaluating the impact of the wireless transmission medium characteristics on the quality of real-time multimedia transmissions.
5. Exploring the capability of different WLAN schemes and topologies to support multimedia traffic.
6. Proposing new mechanisms or strategies to improve the quality of multimedia services distributed over IEEE 802.11 networks.

1.3 Dissertation Organization

In the following, the contents of each of the remaining chapters are outlined.

Chapter 2 “Quality of Experience in Multimedia Services: Evaluation Techniques” presents a global overview of the QoE concept, exploring its different definitions from several points of view. This chapter provides the background necessary

to understand the QoE-perspective given to the rest of the dissertation. Once the multimedia quality assessment paradigm is dissected, the most relevant methodologies for estimating multimedia QoE by means of objective techniques are reviewed. This survey includes methodologies regarding different multimedia services, such as VoIP and Future TV services. Finally, this chapter describes a real experiment for the evaluation of VoIP QoE from both subjective and objective methodologies, exploring the relationship between these approaches.

Chapter 3 “Performance Evaluation of 802.11g Infrastructure Networks Supporting VoIP Traffic: Impact of the PHY Layer and Voice Codec on QoE” discusses about the performance of Wi-Fi networks configured in infrastructure mode, supporting traffic with stringent requirements such as VoIP. To this end, the influence of the wireless physical layer and the voice coding scheme employed on the quality of the voice-calls is evaluated. Additionally, a brief review of the probabilistic propagation models most commonly used by the research community is presented.

Chapter 4 “Performance Evaluation of 802.11g Mobile Ad-hoc Networks Supporting Multimedia Traffic: a QoE Perspective” compares two well-known ad-hoc routing protocols, BATMAN and OLSR, supporting VoIP and video traffic. From the obtained outcomes, the strengths and weakness of these algorithms are discussed. This evaluation is performed in different-nature test-benches, namely, a realistic scenario by using the Emulab platform [9] and a simulation framework such as the discrete-event network simulator Omnet++ [10]. This variety of test-beds permits analyzing several scenarios regarding the traffic pattern employed, the area covered by the network, the number of nodes, or their mobility characteristics, among others. Finally, this chapter presents some novel adjustments to the ad-hoc routing protocols that permit the improvement of the multimedia services’ QoE.

Chapter 5 “Novel Proposals for Improving QoE in 802.11g Mobile Ad-hoc Networks” introduces different approaches developed to improve the QoE and the energy-efficiency of ad-hoc networks supporting multimedia traffic. Specifically, a novel opportunistic routing protocol, so-called JOKER (auto-adJustable Opportunistic acK/timEr-based Routing) is presented. All its features are discussed in detail, showing results comparing its performance against other prominent ad-hoc routing protocols. This study is conducted in both simulation and experimental test-benches, evaluating the implementations of JOKER for these well-differentiated environments. Besides, a

general overview of the opportunistic routing paradigm is given, aiming at presenting the most relevant characteristics of these state-of-art systems.

Chapter 6 “Conclusion” closes the dissertation, remarking the main findings extracted from this thesis. Additionally, a series of open issues regarding the quality of multimedia services over wireless networks and their energy efficiency are summarized.

Fig. 1.1 shows the relationship among the different chapters composing this dissertation.

1.4 Thesis Contributions

The main contributions of this thesis are listed as follows:

Chapter 2: “Quality of Experience in Multimedia Services: Evaluation Techniques”.

- A set of QoE definitions and related background is presented. To this end, two comprehensive surveys, regarding standard and non-standard state-of-art QoE evaluation methodologies for VoIP and Future TV services are provided.
- These surveys present a wide overview of the QoE estimation paradigm and how to apply the best methodology, depending on the scenario and the service under evaluation.
- A real subjective QoE test is presented. The experimental study description dissects the main guidelines to perform this kind of tests, comparing and analyzing the obtained QoE results with those obtained from a QoS perspective.

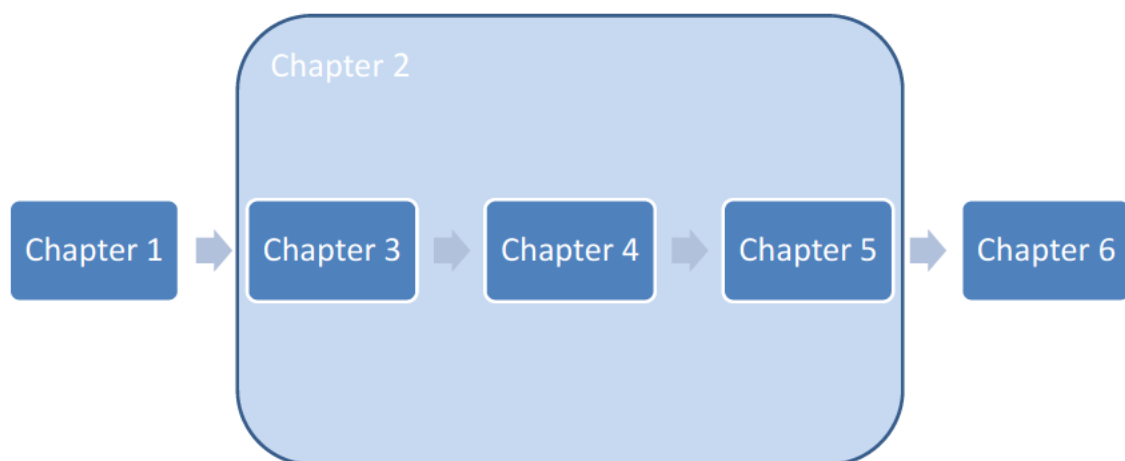


Figure 1.1. Chapters’ organization scheme.

Chapter 3: “Performance Evaluation of 802.11g Infrastructure Networks Supporting VoIP Traffic: Impact of the PHY Layer and Voice Codec on QoE”.

- A complete study from QoE perspective evaluating the capability of IEEE 802.11 networks configured in infrastructure mode to support VoIP traffic is provided.
- A deep characterization of the wireless transmission channel by means of 4 different propagation models is performed.
- The impact of the transmission channel conditions on the QoE of the VoIP service is thoroughly analyzed. Besides, the negative effects of the fading channels on the network's capacity and coverage range are also studied.
- A comparison between the effect of different voice coding schemes (no-compression codec vs. low bit-rate codec) and packetization intervals on the voice calls' quality and the system's capacity is presented.

Chapter 4: “Performance Evaluation of 802.11g Mobile Ad-hoc Networks Supporting Multimedia Traffic: a QoE Perspective”.

- A QoS and QoE exhaustive performance evaluation study about the capability of the ad-hoc routing protocols BATMAN and OLSR to support multimedia traffic (including both VoIP and video-streaming) is conducted.
- Two different test-benches are presented and compared: a simulation platform and a real test-bed.
- A statistical comparison of the results attained in both test-benches that allows studying the correlation between the simulation framework and the realistic test-bed is performed.
- The impact of tuning the BATMAN OGM interval and the OLSR HELLO interval on the quality achieved by the multimedia service flowing through the ad-hoc network is evaluated.
- By using the mentioned time-intervals tuning, some possible enhancements to the BATMAN routing protocol aiming at improving the QoE of the multimedia service flowing through the network are identified and tested.

Chapter 5: “Novel Proposals for Improving QoE and Energy Efficiency in 802.11g Mobile Ad-hoc Networks”.

- Additional improvements for the BATMAN routing protocol are proposed. In this case, a dynamical adaptation of the control-message sending-interval is developed. This enhancement further improves the capability of BATMAN to support highly demanding services, such as video-streaming applications; additionally, it permits reducing the energy consumed by the nodes in routing tasks.

- An exhaustive evaluation of the trade-off between energy consumption and QoE in video transmission with MANET routing protocols (BATMAN, OLSR, and AODV) is conducted.
- An overall overview of the opportunistic routing paradigm is provided.
- A novel opportunistic routing protocol, so-called JOKER, is presented and evaluated.
- A new metric for selecting the best next-hop towards the final destination is developed and tested. This new metric gathers the packet-delivery reliability of the links with the distance-progress towards the final destination.
- Two different candidate coordination algorithms are considered and evaluated.
- Two different JOKER implementations, for simulation and experimental test-benches, are introduced.
- A complete study from QoE perspective, showing the greater capability of JOKER compared with BATMAN to support multimedia services while reducing the energy consumption is presented.

Chapter 2

Quality of Experience in Multimedia Services: Evaluation Techniques

2.1 Introduction

Recently, with the growth of the Internet, a high number of multimedia communication services have been developed. Because of the nature of the human communication, some technical restrictions are imposed when transmitting the media signal: excessive retards, video/audio distortion, etc. are unacceptable. Unfortunately, Internet Protocol (IP), which is the most employed protocol at the network layer, only provides best effort delivery and its service is unreliable. For that reason, different techniques have been proposed in the related literature to provide acceptable levels of Quality of Service (QoS), some of them being successfully deployed by network and service providers, such as Differentiated Services (DiffServ) [11]. Nevertheless, QoS usually focuses only on the objective measure of technical parameters related to network conditions and performance, such as throughput, delay, packet loss, and jitter. Due to the uprising involvement of final users in the forthcoming Internet, for instance in the creation and distribution of multimedia content, an additional way of quality evaluation called Quality of user Experience (QoE) is gaining momentum. In this case, QoE is a measure that focuses on assessing the level of quality the customer perceives when consuming a service. This type of estimation returns a more realistic evaluation of the actual provided quality.

An extended methodology used to evaluate QoE is the Absolute Category Rating (ACR) [2]. This evaluation model outputs a Mean Opinion Score (MOS) which is a subjective rating for the service. For multimedia communications, this score is extracted from a survey in which a large number of participants are asked to assess the quality of a voice call, in a scale from 1 (poor quality) up to 5 (excellent). A MOS test is widely accepted as the most accurate speech quality rating methodology, but its implantation is expensive and does not allow a continuous network monitoring and management. For that reason, several objective models, trying to obtain the greatest correlation in their

results with those obtained by the ACR method, have been recently proposed. The main difficulty of developing a general model to evaluate the QoE of a multimedia service avoiding the human interaction is the subjectivity of the end-user perception, which not always correlates well with QoS measurements [12]. One of the reasons is that human perception is not only influenced by what she is watching or hearing, but also by own human feelings, such as the level of expectation for a particular service, the perception of usability or worthiness of the evaluated service, the cost of the service, or the user's cultural background, among others.

2.2 Quality of Service and Quality of Experience

The ITU (International Telecommunication Union) defines QoS as “the collective effect of service performances, which determine the degree of satisfaction of a user of the service” [13]. To achieve high level of QoS, four related conditions should converge on what it is called the QoS wheel (see Fig. 2.1) [14]: QoS requirements of users, QoS offered/planned by providers, QoS delivered/achieved by the provider, and QoS perceived by the user. The QoS requirements of users are the users' non-technical specification of the level of quality required by the applications. Hereby, users' needs are described in a network-independent language. The QoS offered/planned by the provider represents the level of quality expected to be offered to the user. This specification is usually done in terms of QoS parameters, which derive in a Service Level Agreement (SLA). The QoS delivered/achieved by the provider is a statement of

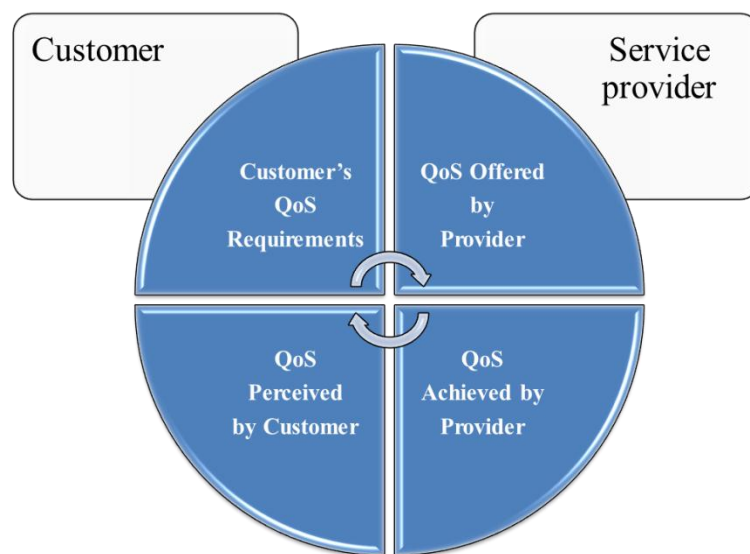


Figure 2.1. QoS wheel (extracted from [14]).

the level of the real quality achieved and delivered to the user. It is given by the current QoS parameters (what should be previously defined) monitored or measured in the network, so that it is possible to compare the planning with the real situation. Observe that this element is the traditional way of evaluating the quality of a provided service, assessing just network parameters such as packet loss, delay, jitter, and throughput. Finally, the QoS perceived by the user is a statement expressing the level of quality that customers believed they have experienced, i.e., QoE. As aforementioned, it is expressed in terms of degree of satisfaction, and usually assessed by customer surveys and comments.

As we can see, the QoS wheel already incorporates QoE as one of its four stages. However, the formal definition of QoE appears later [15] as “the overall acceptability of an application or service, as perceived subjectively by the end-user”. Please note that QoE includes the complete end-to-end system effects (e.g., client, terminal, network, services infrastructures, etc.) and overall acceptability may be influenced by user expectations and context.

Going back to the QoS wheel (Fig. 2.1), it was mentioned that users’ needs may be expressed in a non-technical way, and the specification of the provider’s agreement may come in similar terms. It is clear that there is a straightforward relation among users’ needs and the final users’ perception, the first and the last QoS wheel components. As an initial step, ITU-T Rec. G.1010 [16] identifies QoS categories that are the basis for deriving realistic QoS classes for IP-based services and associated QoS control mechanisms for the underlying transport networks [17], [18]. Although not labeled as a QoE approach, the ITU-T Rec. G.1010 [16] can be seen as one of the first attempts to define the parameters to be used in order to describe performance from the users’ perspective. Note that the user is not concerned about the service implementation, but in the performance that a service presents compared with other similar services from different providers. Following this idea, the ITU-T Rec. G.1010 [16] recommends that performance (i) should be expressed in a way that takes into account all aspects of the service from the customer’s point of view, (ii) focuses on user-perceivable effects rather than their causes within the network, (iii) is independent of the specific network architecture or technology, (iv) can be objectively or subjectively measured at the service access point, (v) can be easily related to network performance parameters, and (vi) can be assured to a user by the service provider. Accordingly, the key parameters

identified in [16] as impacting the user are: delay, delay variation, and loss. But clearly, these parameters are not enough, so other factors affecting the user's overall service experience should be taken into account, e.g., graphical interface, bandwidth consumption, customer support, ease of use, or billing facilities, among others [12], [19]–[23].

For those key parameters selected in [16], [17] identifies their corresponding IP network performance parameters, namely, IP transfer delay (IPTD), IP packet Error Ratio (IPER), or IP packet Loss Ratio (IPLR), among others, and [18] presents the corresponding network performance objectives for each network performance parameters described in [17]. Network performance objectives depicted in [18] are consistent with IETF Differentiation Services standards, and also with ETSI TIPHON 101 329 Part2 definition of speech QoS classes. In addition, the specific ITU-T Recommendation Y.1901 [24] that includes requirements and recommendations for the support of IPTV services also references [18] for QoS performance.

Applications performance requirements are well-understood, but several important contributors are often beyond the network service provider's control (e.g., home networks, local area networks, application gateways, or customer's devices). Some objectives on the performance of customer equipment are available, such as those described in [25] for VoIP terminals and gateways. Combining these objectives with specific network performance levels provided in ITU-T Rec. Y.1541 [18], a view of application performance can be directly related to network performance as agreed by service providers. These objectives are organized in sets called QoS classes that can be matched with well-designed customer's equipment to satisfactorily support various applications. These QoS classes are consistent with both the Rec. G.1000 [14] and G.1010 [16] mentioned previously. The objective values result from analysis of key applications such as conversational telephony, multimedia conferencing, reliable data exchange using TCP, and digital television, in line with network feasibility analysis. The network QoS classes form an important link in the chain of developments required to assure end-to-end performance. Verification that the service meets network objectives is another key of customer interest. Eight network QoS classes compose the basis of agreements between end-users and network service providers, and among network service providers. Applications supported by these QoS classes are: conversational telephony, multimedia conferencing, digital video, and interactive data

TABLE 2.1. QoS classes definition (extracted from [18])

Network Performance Parameter	Nature of Network Performance Objective	QoS Classes					
		0	1	2	3	4	5 (unspecified)
IPDT (ms)	Upper bound of the min IPD	100	400	100	400	1000	U
IPDV (ms)	Upper bound on the $1-10^{-3}$ quantile of IPTD minus the minimum IPTD	50	50	U	U	U	U
IPLR (1×10^{-3})	Upper bound on the packet loss probability	1	1	1	1	1	U
IPER	Upper bound	1×10^{-4}					U

TABLE 2.2. Provisional QoS classes (extracted from [18])

Network Performance Parameter	Nature of Network Performance Objective	QoS Classes	
		6	7
IPDT (ms)	Upper bound of the min IPD	100	100
IPDV (ms)	Upper bound on the $1-10^{-5}$ quantile of IPTD minus the minimum IPTD	50	
IPLR	Upper bound on the packet loss ratio	1×10^{-5}	
IPER	Upper bound	1×10^{-6}	
IPRR	Upper bound	1×10^{-6}	

transfer. Please, see Tables 2.1 to 2.3 for specific QoS quantitative values. Observe that two provisional QoS classes are defined in Table 2.2 whose network performance parameters still need revision based on real operational experience; main difference is that they are more restrictive in terms of packet losses and errors for high rate user applications, for instance for the IPTV service. As extracted from the ITU-T Recommendations mentioned so far, the objectives are primarily applicable if access link speeds are at the T1 (1.544 Mbps) or E1 (2 Mbps) rate and higher, otherwise serialization time of IP packets should be carefully analyzed as it could be the case of some current access network technologies.

As explained so far, much has been advanced in terms of QoS provisioning in telecommunication networks, from standards to network providers' deployment, e.g., Differentiated Services solutions. However, the advance of QoE-aware technologies and services is still in progress. Methodologies focused on measuring QoE are usually classified in two families, namely, subjective and objective approaches. Although subjective tests provide the most accurate results, they are time-consuming, human-dependent, and do not allow the service monitoring in real-time. For that reason, objective QoE-estimation approaches are being proposed by the research community

TABLE 2.3. Matching QoS Classes and Services (extracted from [18])

QoS class	Applications (examples)	Node mechanisms	Network Techniques
0	Real-time, jitter sensitive, high interaction (VoIP, VTC)	Separate queue with preferential servicing, traffic grooming	Constrained routing and distance
1	Real-time, jitter sensitive, interactive (VoIP, VTC)		Less constrained routing and distances
2	Transactions data, highly interactive (Signalling)	Separate queue, drop priority	Constrained routing and distance
3	Transaction data, interactive		Less constrained routing and distances
4	Low loss only (short transactions, bulk data, video streaming)	Long queue, drop priority	Any route/path
5	Traditional applications of default IP networks	Separate queue (lowest priority)	Any route/path

and standardization organizations. In next sections, it is provided a high level description of standard methods related to the QoE assessment of multimedia services, focusing on how to apply them in VoIP and Future TV services. The state of the art of non-standard proposals is also explored by presenting a comparison of approaches.

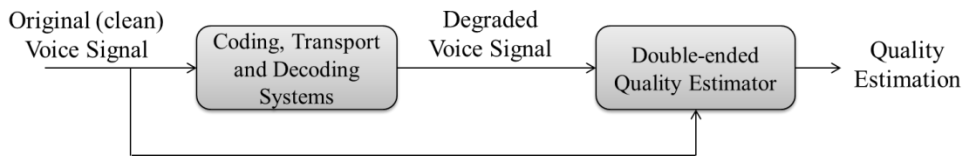
2.3 VoIP QoE Estimation Methodologies

VoIP is one of the most promising services for the end-users. It provides low-cost calls between communicants located at any place around the world. However, the penetration of this service has been slower than that expected some time ago. The main reason for this low adoption of VoIP services is the end-users' perception about the quality of the calls [26], which does not drive them to abandon the traditional telephone service. For that reason, all stakeholders, namely, Telco operators, service providers, and VoIP equipment manufacturers, face the challenge of improving the quality provided to end-users. The first step to understand customers is evaluating the service quality from their point of view (QoE). VoIP quality can be measured in both, listening (MOS-LQ, Listening Quality) and conversational terms (MOS-CQ, Conversational Quality). The former case takes into account the audio fidelity, distortion, speech, and noise levels in one direction, so it deals with how users rate what they hear during a call. The latter takes also into consideration bidirectional effects such as delay or echo, so it refers to how users score the overall quality of the call based on listening quality and their conversational experience during a call. As mentioned above, subjective tests are being replaced by objective methodologies to accomplish an automatic assessment of

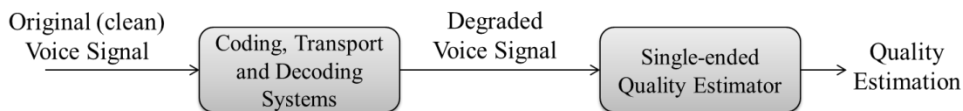
the quality provided to the customers in real-time. These methodologies can be broadly classified into intrusive and non-intrusive models (see Fig. 2.2).

On one hand, intrusive models (Fig. 2.2 (a)), also known as comparison-based, full-reference, or double-ended models, use the original speech-signal to compare it with the degraded one, when it arrives to its destination. The problem of this procedure is that the source information or original voice is needed at the destination extreme, which in most cases is impossible since users are physically apart. Furthermore, it is not possible to use intrusive methods in real-time, since the full reference-signal is needed. Perceptual Evaluation of Speech Quality (PESQ) [27], Perceptual Objective Listening Quality Assessment (POLQA) [28], and Enhanced Modified Bark Spectral Distortion (EMBSD) [29] are examples of these models.

On the other hand, non-intrusive models (Fig 2.2 (b)), sometimes called single-ended or output-based methods, estimate the service quality by assessing objective parameters at any point of the VoIP communication path, without needing a reference signal. These methods allow detecting bottlenecks and also permit intermediate devices (e.g., routers or switches) and higher- level applications to generate real-time alarms when a decrease in the quality of the service is detected. Thus, real-time concept reaches a remarkable relevance because it allows to take decisions to improve the levels of quality along the duration of the call, so the final user increases her perceived QoE before the call ends [30], [31]. In turn, non-intrusive models can be classified as signal-based models, parametric models, and packet-layer models. The former estimates the QoE by processing directly the human speech, analyzing the distortion introduced in the voice signal; an example of this methodology is the ITU-T standard P.563 [32].



(a) Intrusive quality model



(b) Non-Intrusive quality model

Figure 2.2. Objective quality estimation methodologies.

On the other hand, parametric models, such as ITU-T Rec. G.107 (E-model) [33], base their QoE estimations on the assessment of different impairments introduced by the network and the encoding schemes. Parametric models need a full knowledge of the system, from end-terminals to network equipment and data-links employed in the VoIP transmission in order to tabulate the different sources of impairments that could affect the communication. Finally, the packet-layer models only employ the information which can be extracted from the different headers on the multimedia packet, e.g., IP/UDP/RTP. The parameters analyzed by these models to obtain their quality estimations are packet loss, delay, coding scheme, among others. An example of this methodology is the ITU-T Rec. P.564 [34].

Quality measurements methods have notably evolved in telecommunication networks [3], [35]–[37]. It is worthy to remark the work by Jelassi *et al.* [3] in which authors presented a complete overview of existing subjective and objective methodologies for the assessment of the QoE of voice conversations as well as some practical procedures for measuring VoIP QoE. Other works [35], [36] also reviewed quality estimation methodologies but from a more general perspective, compared to the detailed algorithm descriptions that are provided in this work. Finally, the study by Mahdi and Picovici [37] provided a wide overview of objective quality assessment models, but with less emphasis on non-intrusive proposals than this work. It is considered that the capability provided by the single-ended models consisting on taking measurements at real-time, is the key for the future development of high-quality VoIP services which could react to QoE drops during the call. Compared with all these works, it is also provided a more extensive review about the different author's proposals which, in some cases, suppose a great improvement to the standard models.

In the following sections, it is presented a complete review of current objective non-intrusive models for VoIP QoE evaluation, analyzing their strengths and weaknesses. Some clues related to the best suited scenarios for the different reviewed approaches are also provided.

2.3.1 Standard Models

2.3.1.1 E-Model

The E-model [33], originally designed as a telecommunication transmission planning tool, has become one of the most popular methods to evaluate the quality of a voice

transmission system. This parametric model takes into account several tabulated transmission impairments, such as delay, echo, codec distortion, etc. measured by In-Service, Non-Intrusive Measurement Devices (INMDs) [38] (a deep description of this devices is provided in section 2.4.3). Impairments assessed by INMDs form part of an additive rating scale, called R , which estimates the conversational quality (MOS-CQE, Conversational Quality, Estimated) of a voice call. R can be directly mapped to a MOS scale, so it is also employed to predict customer's conversational QoE (MOS-CQS) (see Table 2.4). As described in ITU-T Rec. G.107 [33], R is obtained through expression (2.1),

$$R = R_o - I_s - I_d - I_{e-eff} + A \quad (2.1)$$

where R_o is the basic signal-to-noise ratio, including background noise and transmission noise; I_s represents impairments that occur simultaneously with the voice signal, such as a non-optimal side-tone level; I_d includes the impairments caused by delay; I_{e-eff} denotes the impairments due to low bit-rate audio codecs and randomly distributed packet losses, and A is the advantage factor, which allows for compensation of impairment factors when the user benefits from other types of access advantages (e.g., a mobile environment). As shown in Table 2.4, R ranges from 0 (lowest possible quality) to 100 (optimum quality).

The R to MOS transformation is made by using the expression (2.2)

$$\begin{aligned} R < 0: MOS &= 1 \\ 0 \leq R \leq 100: MOS &= 1 + 0.035 R + 7 \cdot 10^{-6} R(R-60)(100-R) \\ R > 100: MOS &= 4.5 \end{aligned} \quad (2.2)$$

The calculations to obtain the aforementioned impairments involve many input parameters and complex mathematical formula; for those reasons, the ITU-T proposed a reduce E-model [39], taking into account just the impairments related to the

TABLE 2.4. R to MOS mapping

R	User Satisfied	MOS
90 - 100	Very Satisfied	4.34 – 4.50
80 - 90	Satisfied	4.03 – 4.34
70 - 80	Some Users Dissatisfied	3.60 – 4.03
60 - 70	Many Users Dissatisfied	3.10 – 3.60
50 - 60	Nearly All Users Dissatisfied	2.58 – 3.10
0 - 50	Not Recommended	1 – 2.58

transmission over the network, and setting the rest of parameters to their default values (see Table 2 in ITU-T Rec. G.107 [33]). Using this simplified model, the expression for R is reduced as shown in (2.3),

$$R = R_o - I_{dd} - I_{e-eff} \quad (2.3)$$

where I_{dd} represents the impairments due to transmission delay in echo-free connections, being calculated as a function of the absolute transmission delay. I_{e-eff} includes the same effects as in the original formula taking into account that, for each codec, this impairment is calculated as a function of packet loss probability.

The description given in ITU-T Rec. G.107 [33] presents the E-model as a planning tool for narrow band transmissions (300–3400 Hz), but an extended version for wideband (50–7000 Hz) has been recently presented [40]. Although some wideband E-model predictions are currently under study, this extended model captures the effects of several factors ignored in first versions, such as low-bitrate wideband coding or VoIP packet loss degradations.

Based on this transformation [41], the maximum value for R (wideband) is now 129; thus, the quality improvement introduced by wideband transmission is quantified on the R -scale. Expression (2.1) applied to wideband transmission undergoes some changes as following; the impairment parameter I_s and the advantage factor A are not adequately analyzed yet, so both take a value equal to 0. I_d takes into account talker/listener echo, in addition to the absolute transmission delay and, finally, I_{e-eff} includes the effect of random or bursty packet loss and speech coding. An extended expression for R factor in wideband is shown in (2.4),

$$R_{WB} = R_{o,WB} - \{I_{dte,WB} - I_{dle,WB} - I_{dd}\} - \left\{ (I_{e,WB} + (95 - I_{e,WB}) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}}) \right\} \quad (2.4)$$

where $R_{o,WB}$ is the basic wideband signal-to-noise ratio, including noise sources such as circuit noise and room noise; $I_{dte,WB}$ and $I_{dle,WB}$ give an estimate for the impairments due to talker/listener echo; I_{dd} includes the impairments caused by delay; $I_{e,WB}$ and B_{pl} are codec-specific values and represent the impairment due to low bit-rate coding and the packet-loss robustness factor, respectively; values for these parameters can be found in Appendix IV of the ITU-T Rec. G.113 [42]. P_{pl} is the packet-loss probability and $BurstR$ is the burst ratio, which represents the bursty pattern of the transmission.

An additional extension of the wideband E-model was also presented in [43]. In this work, the E-model rating scale is extended to an R maximum value of 179, aiming to quantify the quality gain obtained by using super-wideband speech transmission. This extension corresponds to a quality gain of 39% as compared to wideband, and 79% as compared to a narrowband channel. As discussed above, wideband and super-wideband E-models are currently under development and they are subject to changes, especially parameters $I_{e, WB}$ and B_{pl} ; consequently, the accuracy of quality predictions is not guaranteed, yet.

2.3.1.2 Improvements to the E-Model

The E-model has a wide related bibliography and it has been studied and re-implemented by several authors, adding improvements to impairment calculations and characterising the model to different environments.

Conway [44] presented a methodology employing the ITU-T standard PESQ with the aim of calculating the value of I_e for a finished communication. Given that PESQ is an intrusive method that needs a reference signal, Conway proposed to extract the observed frame erasure process of the call and apply it to a clean signal. Thus, a clean reference and its pseudo-received (degraded) version are available in the receiver; the quality estimation obtained by PESQ employing these artificial signals is an approximation to the quality score that would be obtained by this method applied directly to the real voice signal and its transmitted version. Once the MOS score is calculated, it is mapped to the R score, and the equipment factor I_e is calculated as, $I_e = R_o - R$, where R_o is the basic signal-to-noise ratio, which is assumed to its default value as defined in the ITU-T Rec. G.107. This methodology has some important drawbacks such as the assumption that the main perturbation affecting speech quality is frame erasures, ignoring the effect of several perturbations such as attenuation, noise, or echo. Another issue of this model is that its accuracy lies on the capability of the adopted reference signal to capture the characteristics of the original speech signal; for those reasons, the possible target scenarios are very limited.

Clark [45] proposed a method based on the Gilbert model to include the effect of bursty packet loss and recency, developing a time-varying listening quality measurement algorithm. Since packet loss and delay are not constant values during the call, the author characterized I_e and I_d as instantaneous impairments, calculating the final

R for the call averaging the different values of these terms during the connection. This average is assessed taking into account the “recency effect” on the users, i.e., the quality perceived at the end of the call has a greater weight in the final subjective quality score than the quality perceived at the beginning of the communication. This model, so-called VQMon, was proposed as P.564 candidate algorithm [34]; this fact will be detailed in section 2.3.1.6. Another work supporting the “recency effect” was presented by Rosenbluth [46], in which the author already obtained similar results to those attained by Clark, demonstrating that the last part of the call had a bigger importance in the final perceived QoE. However, Gray *et al.* found the opposite in a different study [47], where they showed that the first part of a speech signal has the greatest weight in the overall quality perception; this effect was termed as “primary effect”. In the light of the results obtained in these works [45]–[47], additional studies would be necessary in order to determine the existence of either recency or primary effects.

An implementation of the algorithm described in [45] can be found in the work by Carvalho *et al.* [48], in which it was evaluated the MOS variations during several simultaneous VoIP transmissions in a real scenario. These authors also proposed corrections in R to MOS conversion, fixing out-of-scale MOS values (below 1) when R takes values under 6.5. In addition, I_e miscalculation was also fixed, obtaining this parameter separately for burst and gap periods. Although the results showed reasonable correlation with user ranking, some concerns arose: (i) the subjective comparison could not be taken as final since the subjective tests were not carried out following standard recommendations (e.g., the number of participants was too low), (ii) it was not possible to identify the exact effects of lost packets during talk-spurts, and (iii) additional work would be recommendable to incorporate smart packet loss events which would be generated by the codecs.

Cole and Rosenbluth [49] made use of the reduced version of the E-model, emphasizing the accurate estimation of I_d and I_e impairments. The rest of parameters, not related to the transport network, such as the quantizing distortion or room and circuit noises, were set to their default values. In this work, I_d was calculated as a function of the ear-to-mouth delay, which has three main contributions: codification, de-jitter buffer, and network delays. For the calculation of I_e authors established two main contributions, namely, packets lost in the network and packets dropped by the de-jitter buffer. Results of this work showed that R is very sensitive to packet loss, taking also

into account the aforementioned different sources of delay. However, this algorithm does not take into account several impairments inherent to VoIP, such as bursty packet loss behaviour, distortion introduced by speech coding scheme, or variations introduced by the de-jittering buffer. Another limitation of this model is that R is calculated using experimental parameters, validated for very specific conditions of codec, packet size and error pattern; a generalisation of these results would be recommendable and necessary.

Another work that investigated the effect of packet loss and jitter in speech quality was presented by Ding and Goubran [50]. Authors modelled the packet loss impairment, I_e , as a logarithmic function of the packet loss ratio; in turn, the effects of the jitter were quantified by adding a new impairment, I_j , to the R expression. This new parameter I_j was estimated as a function of the size of the de-jitter buffer and the Hurst parameter. Authors obtained an accurate quality prediction when packet loss was below 10%. In turn, Ren *et al.* [51] also made use of the jitter impairment, I_j . They analyzed the impact of the jitter standard deviation on the MOS score by assessing delay variations under different packet loss and de-jittering buffer size conditions. Results obtained showed good accuracy with PESQ predictions. In both models, additional work would be needed to assimilate these results in more heterogeneous network conditions (e.g., different coding schemes) aiming to assure the generalization capability of the algorithms.

Zhang *et al.* [52] proposed an E-model enhancement focused on the measurement of packet loss burstiness. This improvement denoted as Q-model, analyzes the packet loss pattern with a low computational-complexity algorithm and included a burstiness parameter in the calculations for the packet loss probability. This factor plays a main role in the estimation of the equipment impairment, I_e . This new burstiness parameter may be taken as an enhancement to the aforementioned factor $BurstR$, from the original E-model. The burstiness parameter captures the local burstiness of dropped packets using an efficient window-based algorithm, i.e., analysing the packet loss pattern of previous packets when a dropped packet is detected. The average value for P_{pl} is estimated as the sum of the measured P_{pl} value and an additive weighted combination of gathered values for the burstiness parameter. Authors claimed that improved measurement accuracy was obtained under bursty packet loss conditions when compared to the standard E-model. Nevertheless, the parameters in the model would

need to be refined for different codecs and Packet Loss Concealment (PLC) algorithms, because the results under heavy burst packet losses were not completely satisfactory.

From a different perspective, Grah and Radcliffe [53] questioned the applicability of the E-model to VoIP transmissions, as it was designed for systems that worked in different scenarios to those currently used for VoIP. A new expression for R , called VoIP E-model (VoIP-EM), was proposed. VoIP-EM includes the relevant impairments for current real-time systems, such as coding scheme, packet loss and delay, ignoring those impairments whose impact could be dismissed in VoIP quality, as echo, quantization, and loudness impairments. Thus, the expression proposed for VoIP-EM is shown in (2.5),

$$VoIPeM = N_o - SE \quad (2.5)$$

where N_o is the default initial quality value according to the network setup and codec type and SE represents the sum of all errors, which are calculated as a linear or non-linear function of the aforementioned delay, packet loss, and codification impairments. However, this work did not provide any analytical result or performance evaluation, in order to show the accuracy of the algorithm. Specific configuration parameter values were not included either, so the efficiency or generalisation capability of this proposal was not manifested.

A more complete work, in which the E-model was also redefined to work in packet-switched networks, was presented by Meddahi and Afifi [54]. Authors assumed that the main factor affecting the speech quality in VoIP environments are: Analog/Digital (A/D) and Digital/Analog (D/A) conversions, coding algorithm, bandwidth, jitter, delay, and packet loss. By analysing and adapting these elements to the classical parameters in the R additive scale, a derived model for datagram transport, called Packet-E-model (P-E-model), was obtained. The new expression to estimate R is shown in (2.6).

$$R_p = R_{op} - I_{dp} - I_{ep} + A \quad (2.6)$$

Observe the similarity of this formula with the reduced E-model expression (2.3). In this case, R_{op} includes a new factor to represent the effect of packet switched noise, whereas I_{dp} and I_{ep} are obtained as indicated in the standard ITU-T Rec. G.107. R_p is calculated for every VoIP packet and it may be averaged over a longer period, e.g., during a phone call. Authors tested the P-E-model in both simulated and real test-beds.

Results showed the evolution of the estimated MOS under several packet loss and delay conditions, evidencing good model response against impairment variations.

Lewcio and Waltermann [55] noted that the standard E-model is not capable to predict the quality where narrowband and wideband codecs are present in the same transmission. Next Generation Networks (NGN) will be a mix of diverse VoIP and Internet access providers, using both, narrowband and wideband codecs, equipment, etc. Thus, a VoIP communication will suffer several wideband-narrowband switches, with the consequent effect on the speech signal quality. The result of not taking this issue into consideration is an overestimation of the call quality. For that reason, authors introduced a new codec impairment factor, I_{esw} , which quantifies the effect of speech bandwidth switching during active VoIP sessions. The codec switching impairment factor is calculated as a function of the difference of quality between the codecs involved, no matter bandwidth switching direction. Results showed that the improved model has better accuracy than the original one. Further tests would be needed because just two experiments were carried out to validate this enhancement; additionally, some empiric key parameters were calculated out of these few experiments, so it would be interesting to evaluate the generalization capability of this model.

Assuming that VoIP frames can be classified as speech or silence frames and that loss on silence periods do not affect negatively to speech quality, Li *et al.* [56] presented an improvement to the E-model using an active packet detection algorithm. By detecting the category for each lost packet, dropped silence frames are not taken into account for assessing P_{pl} , including just the lost speech packets. Adding this enhancement into the E-model, authors obtained a better MOS estimation accuracy compared to the original method.

Works reviewed so far (except the P-E-model [54]) focused on wired environments. However, due to the heterogeneous nature of current and future Internet, the performance of the E-model in wireless networks is also of high interest. Thus, Falk and Chan [57] investigated the impact of wireless-VoIP degradation on the performance of the E-model, among other standards such as P.563 and PESQ. Through factorial analysis of variance tests, authors found that the performance of the aforementioned algorithms is sensitive to several degradation sources such as noise level and codec-PLC type. Focusing on the E-model, they also suggested several significant two-way interaction effects, such as codec and noise type or codec and noise level, concluding

that the E-model accuracy is limited for wireless-VoIP scenarios. Work by Picovici and Nelson [58] focuses on the main impairment introduced by the wireless networks, i.e., the variability of packet loss. Authors added a new parameter, I_p , into the R scale with the aim of accounting for the perceptual relevance of the packet loss variations. I_p is calculated by computing the Euclidean-based median distance between a vector representing the perceptual features of the clean signal and another including the speech features of the received signal. An important drawback is found in this proposal, regarding the needing of the clean signal, which transforms this model into a reference-based methodology.

Another work considering also the impact of wireless-system issues, such as temporal disconnection, was presented in [59]. This model, so-called PEVOM (Perceptual Evaluation of Voice over MANETs), detects the loss of nodes connectivity and estimates separately the quality experienced in connected and disconnected periods. When users are connected, the instantaneous transmission quality is calculated using the reduced version of the E-model [39]. On the other hand, when users loose connectivity, PEVOM predicts the instantaneous transmission quality employing a formula taking into account the users' dissatisfaction for the issues during the call. The performance of PEVOM was compared with that attained by VQMON in ad-hoc scenarios, obtaining higher level of accuracy on the predicted quality by PEVOM. According to the authors, VQMON underestimates the final quality of the communication when disconnections happen during the call; however, no other comparative results (e.g., subjective test, or PESQ) were showed to confirm this hypothesis.

An extension of this work considering the impact of different loss distributions was proposed by Jelassi *et al.* [60]. This algorithm, denoted as NIDA (Non-Intrusive Disconnection Aware assessment algorithm) takes also into account the impact on the QoE of the intermittent loss of connectivity in wireless scenarios. To this end, a new wireless channel model was developed, based on a continuous-time Markov model. Authors modeled the stochastic process with 3 states: I (isolated), B (burst), and D (disconnected), evaluating the QoE separately for each state. At the end of the call, all these measurements are combined to predict the overall quality of the communications. The results for NIDA, show a big correlation with the PESQ model.

Finally, and from a different perspective, Passito *et al.* [61], [62] presented a tool, based on the E-model, to quantify the impact of using the IP Security protocol (IPSec)

in an IEEE 802.11 wireless local area network. Authors also made use of an Ethernet wired environment to compare the results obtained. In both wireless and wired scenarios, a decrease of MOS values was observed when security algorithms were employed; additionally, the E-model-based QoS tool described in that work was employed to manage the network security parameters in order to reach higher levels of quality in the VoIP communications.

2.3.1.3 Call Clarity Index

The Call Clarity Index (CCI) is a parametric estimation model [63] that combines the measurements made by INMDs [38] into a single index correlated with the call quality score determined in subjective tests (e.g., MOS).

INMDs are devices used for both, the assessment of voice-grade parameters (such as speech level, noise level, echo loss, and speech echo path delay) and the evaluation of parameters associated with digital transmission systems. INMDs can be used in circuit and packet switched networks, with the aim of assessing the impairments introduced on the voice-grade channels being transported (e.g., packet loss ratio, delay, and jitter) [38]. Thus, CCI method predicts the conversational quality (MOS-CQO) from the objective parameters measured by the INMDs. The output from the CCI is a score ranging from 1 (bad quality) to 5 (excellent quality) according to the MOS scale described in previous sections.

It is worthwhile to mention that the data produced by INMDs present certain issues. One drawback is due to the INMDs location in the middle of the communication path; so that, values generated by these devices do not represent what the user actually perceives. This issue is overcome by incorporating a more complete model of the network into the CCI algorithm. Another problem related to the INMDs is the threshold imposed for every parameter measured; once this threshold is reached, the device sets this parameter as unacceptable and reports a problem, without taking into account the relationship between different parameters, which could combine each other and cancel the negative effect of one of them. CCI overcomes this problem by using a human model perception that combines the effects of the different input parameters [64]. For VoIP quality evaluation, CCI receives as input the following measures from the INMD: speech level, noise level, echo loss, echo path delay, IP packet loss ratio, and delay variation, in both directions. Then, these parameters are processed through the series of

blocks shown in Fig. 2.3. The assumptions-model block adds the aforementioned information required by the CCI algorithm about the network that is not available from the INMD measurements. Without this information it is impossible to produce a complete description of the end-to-end network. Setting the INMD in the middle of the transmission path, the system is described in terms of two parameters, namely, sending and receiving sensitivity. The former describes the frequency response of the talker's telephone and line to the INMD. The latter includes the frequency response of the listener's telephone and line to the INMD. These assumptions are combined and generate a complete description of the network which is passed to the next stage. The loss and noise perception model accounts for the frequency selectivity of the human ear and the noise masking effects on the connection; side-tone and room noise effects are accounted, too. This block calculates a conversational speech quality prediction for each listener which is passed to the next block.

Finally, the echo and delay perception model modifies the previous-stage result accounting for delay and echo present on the connection by evaluating network parameters. The results of this last step are the CCI outputs which are two scores, one for each end of the connection, i.e., the MOS estimation for both directions of the communication. These are independent results, so they should be analyzed and reported separately.

Fig. 2.4 depicts the correlation between the results obtained by the CCI method and subjective tests reported [64]. Although the results in Fig. 2.4 show good accuracy for the CCI model, it seems that very low number of experiments was conducted, namely, 10. So that, the accuracy of CCI predicting subjective MOS should be further investigated and additional analysis with different network conditions, codecs, etc. is needed. The only INMD implementation found in the literature corresponds to the work by Hartanto [65]. In this study, the author explained the qualitative characteristics of the

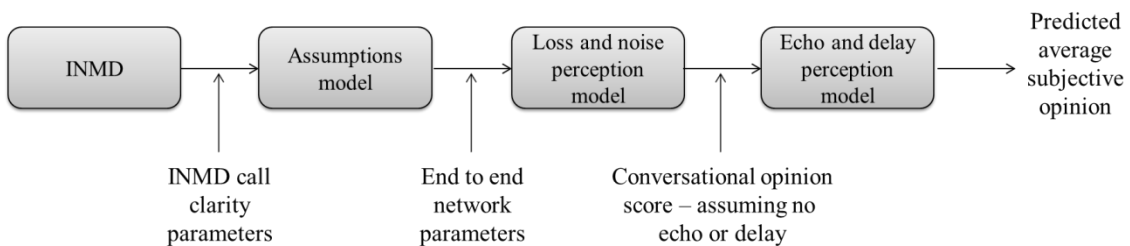


Figure 2.3. CCI blocks (extracted from [64]).

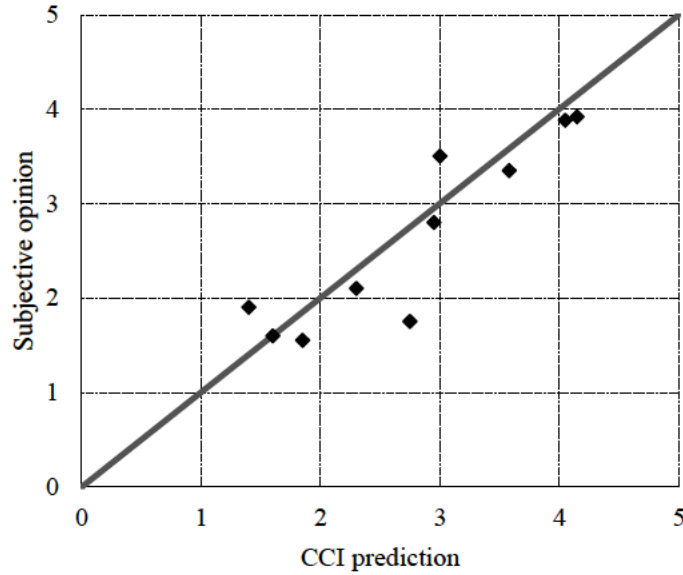


Figure 2.4. CCI accuracy versus subjective tests (extracted from [64]).

proposed tool, so-called QualiView, and gave an overview of the principal advantages of INMDs versus intrusive models. The main purpose of QualiView is to provide a means to ascertain overall transmission link quality in order to (i) isolate problem transmission parameters, (ii) understand how the network is utilized and its influence on the communication quality, and (iii) help eliminate switching problems.

2.3.1.4 P.563

This ITU-T Recommendation [32] describes a single-ended, signal-based method for objective speech quality assessment in narrowband telephony applications. This methodology was selected based on the results of a competition held by the ITU-T from 2002 to 2004. SEAM (Single Ended Assessment Model), a method resulted from the collaboration of SwissQual's NiNA (Non-INtrusive speech quality Assessment) [66], Psytechnics' NiQA (Non-Intrusive speech Quality Assessment) [67], and Opticom's P3SQM (Perceptual Single Side Speech Quality Measure) [68] algorithms, narrowly beat the other proposal, ANIQUE [69], hence being adopted as the ITU-T P.563 recommendation; an extension of the latter method, ANIQUE+ [70], became an American National Standard Institute (ANSI) standard in 2006 as it will be discussed in section 2.3.1.5.

P.563 is able to predict the listening quality (MOS-LQO) in a perception-based scale taking into account the full range of distortions occurring in public switched telephone networks; therefore, this model allows real-time measurements to estimate the MOS of

a voice call at any point of the path between users. The P.563 model is formed by the series of blocks shown in Fig. 2.5. As explained in the work by Malfait *et al.* [71], the first step consists in the pre-processing of the received signal. This stage begins with a speech level adjustment; after this normalization step, the pre-processing block generates two additional versions of the degraded signal; one of them is filtered, using a response similar to the properties of the modified Intermediate Reference System (IRS) frequency characteristics described in the ITU-T Rec. P.830 [72], and the other one is filtered using a fourth-order Butterworth high-pass filter. This filtering is carried out to simulate the frequency response of the telephone terminals.

Finally, a Voice Activity Detector (VAD) is used to separate voice and non-voice parts. Once the signal is pre-processed, it passes through different blocks that generate different quality estimations. The first block uses the vocal track model introduced by Gray [73] in order to extract parameters from the distorted speech, which allows measuring the unnaturalness of the speech and, therefore, its quality degradation. The second block uses a full reference perceptual model for estimating the quality of the degraded speech. As discussed above, this type of methods needs a non-distorted reference signal to measure the levels of quality of the transmitted signal. Since this clean signal is unavailable at any point of the transmission path, a speech reconstruction algorithm is employed. This process erases noise and other distortions from the speech signal and generates a quasi-clean signal, which is employed as reference by the speech quality measurement model. The quality estimation is carried out by a model based on the PESQ model [27]. Finally, the last block focuses on the detection of particular degradations such as noise estimation or temporal clipping and robotization detection.

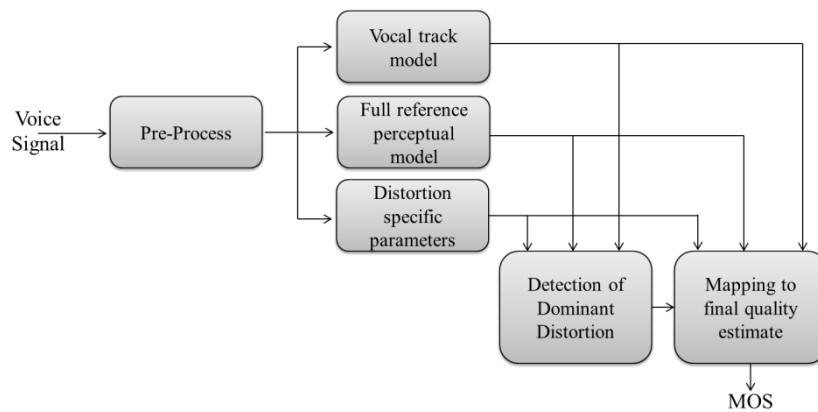


Figure 2.5. P.563 blocks (extracted from [71]).

Once this complex analysis is done, a dominant distortion class is determined and a class-specific subset of the extracted parameters is used to estimate signal quality. This intermediate rate is linearly combined with other eleven parameters to derive a final listening quality score; as discussed above, this score can be mapped to the subjective MOS by an almost linear regression. The estimations accuracy attained by this model was presented in the ITU-T Rec. P.563 [32], by assessing the correlation obtained with subjective tests, employing a number of speech-signal databases. The Pearson correlation coefficient obtained by employing P.563 was 0.88, a little lower than that reached by PESQ for the same test bench, 0.93. However, the advantages of employing a non-intrusive algorithm overcome this loss of accuracy in the MOS prediction.

P.563 counts with very few implementations on the literature compared with the E-model. A simplification of the quality mapping function of P.563 was proposed by Ekman and Kleijn [74]. Authors proposed a method to create a mapping from the features of P.563 to a single objective quality score using non-negative matrix factorization, obtaining an improvement on the algorithm performance when speech databases containing noise suppression data are processed.

Falk *et al.* [75] presented an in-depth analysis of the P.563 algorithm, proposing some modifications to improve its accuracy when synthesized speech data is used, since P.563 was designed to work with natural speech. In particular, authors made use of a regression tree in order to map the features computed by the P.563 algorithm into a final quality rating, which showed higher correlation with subjective results than the original model, for synthesized speech data.

Finally, Abareghi *et al.* [76] introduced an enhancement to the P.563 model to adapt its features to VoIP conditions. Authors declared mute-length, sharp-decline, and speech-interruptions as the most sensitive parameters to network variations. Consequently, a new distortion class and a priority for this new class are defined using those parameters. This new class is added to the P.563 quality estimator. However, this work did not provide any comparison between the accuracy obtained by the standard and the proposed model, so additional work would be needed in order to demonstrate the validity of this enhancement.

2.3.1.5 ANIQUE+

The Auditory Non-Intrusive Quality Estimation Plus (ANIQUE+) model is an ANSI (American National Standards Institute) standard for non-intrusive, signal-based, estimation of narrowband speech quality. ANIQUE+ estimates the listening quality (MOS-LQO) of a voice call based on the functional roles of human auditory systems and the characteristics of human articulation systems [70]. Therefore, the output of this algorithm is an estimation of the listening quality that could be obtained by subjective auditory tests like MOS.

The ANIQUE+ algorithm focuses on measuring the overall distortion affecting the voice signal and mapping this distortion to a MOS value. Fig. 2.6 shows the block structure of the ANIQUE+ model. In the first block, the input is pre-processed, normalising the level of the speech signal and applying the Receive-side Modified Intermediate Reference System (RXMIRS) receive filter in order to reflect the frequency features of the handset used in ITU-T listening tests. This pre-processed speech signal is passed to the next step where other types of distortion are analyzed. The ANIQUE+ algorithm uses three distortion measurement modules. As described in Kim and Tarraf's work [70], the articulation analysis block separates the original speech signal into time frames and the perceptual distortion to estimate the overall distortion of

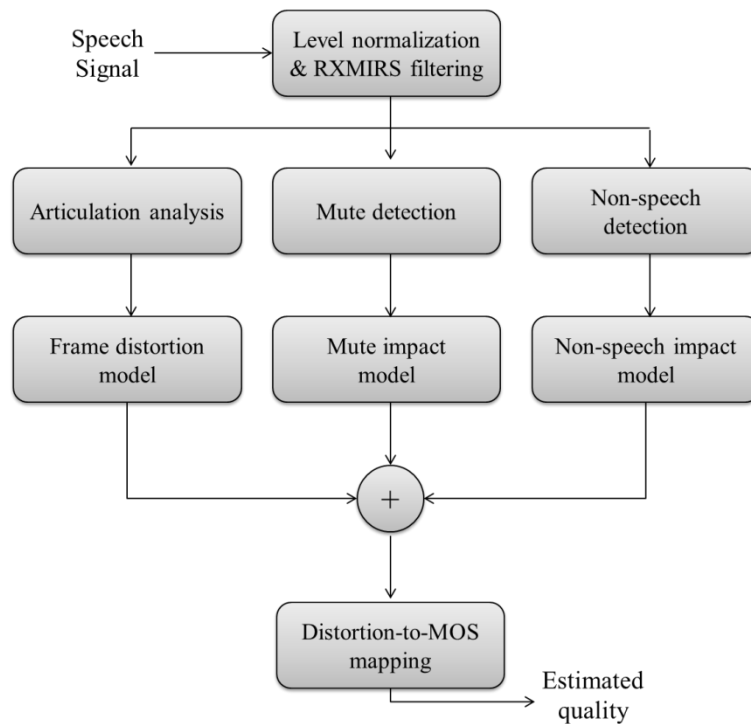


Figure 2.6. ANIQUE+ blocks diagram (extracted from [70]).

the input speech signal. The mute detection module detects unnatural mutes in speech signals, obtaining a speech activity profile. These measurements are passed to the mute impact module that estimates the impact on the speech quality degradation of the unnatural mutes in the speech signals; for this analysis, the recency effect [45] is also taken into account. Finally, the non-speech modules detect the effects of annoying non-speech activity and quantify its impact on speech quality. Kim and Tarraf [70] presented results in which ANIQUE+ overcame P.563's accuracy, reaching very close performance to the intrusive PESQ model. Particularly, making use of 10 non-training databases, ANIQUE+ reached an accuracy of the 93.9% while P.563 only reached the 84.8% of performance. However, the best results were still obtained by PESQ, attaining a 95.3% of accuracy.

2.3.1.6 P.564

All the measurement methods mentioned so far, except the E-model, evaluate the quality of the VoIP service by analyzing speech parameters, e.g., SNR, echo, silent periods, coding distortion, etc. Thereby, aforementioned algorithms need to de-packetize vocal signal contained in IP flows to evaluate the speech payload. This task requires a considerable computational processing, compared with parametric models. For that reason, ITU-T Study Group 12 (SG12) "Performance, QoS and QoE" was concerned about the development of a model that could estimate subjective listening quality (MOS-LQO) based on information obtained just from packet headers. Consequently, the SG12 launched the P.VTQ (Voice Transmission Quality) competition during the period 2002-2004. Two different methodologies were presented: Telchemy's VQMon [45] (previously discussed) and Psytechnics' PsyVoIP [77].

PsyVoIP makes use of the block structure shown in Fig. 2.7. When a call stream is detected, PsyVoIP assigns a specific ID to it, and pre-processes the signal, discarding the useless fields of the headers. Then, packets are realigned and a VAD algorithm is used to mark them with a "voice" or "silence" flag. This process yields to more accurate results, since distortion in silence packets has lower impact on quality than impairments on voice packets. Finally, the quality estimation is calculated as a function of network parameters extracted from the processed stream. It is also remarkable that this model takes into account the particular features of the different manufactured edge-devices, such as VoIP phones or gateways, by using calibrated formulas and weighting



Figure 2.7. PsyVoIP blocks diagram (extracted from [77]).

coefficients to each specific device. No winner for the VTQ quest was selected, but this competition served to develop the Recommendation P.564, which specifies the minimum criteria for objective speech quality assessment models that predict the impact of observed IP network impairments on the one-way listening quality experienced by the end-user in IP/UDP/RTP-based telephony applications [34]. Originally specific to narrowband (3.1 KHz), the P.564 Recommendation already includes an extension for wideband (7 KHz) telephony.

To the author's knowledge, there is still no model elected as standard for this recommendation. The candidate models compliant with this Recommendation should produce a MOS estimation as defined in the ITU-T Rec. P.800 [2]. This evaluation should be done ignoring the voice payload; thus, conversational impairments such as speech level, background noise, side-tone level, or echo have to be disregarded. On the other hand, the voice codec used should be always taken into account. Finally, a model compliant with this Recommendation should be able to be deployed in endpoint locations as embedded monitoring agents, at mid-network monitoring locations, or a combination of both.

The accuracy criteria of candidate models will be based on a comparison of the model's performance with the intrusive PESQ algorithm using the output mapping defined in the ITU-T Rec. P.862.1 [78]. Following this criteria, Neves *et al.* [79] proposed a model based in the E-model, calibrating its parameters, such as *BurstR*, by a regression analysis, which modelled the relationship between the MOS obtained by PESQ and that attained by their algorithm. The proposed model was validated according to ITU-T Rec. P.564 requirements and the results fit in standard accuracy class C2 defined in this ITU-T Recommendation.

2.3.2 Non-Standard Models

In addition to the standard models reviewed in the previous section, there are an increasing number of algorithms proposed by several authors aiming to improve the accuracy of standards by using different techniques. In the following, the most relevant

techniques are discussed. For simplicity, the non-standardized proposals have been classified into the following categories: based on Gaussian mixture models, based on neural networks, based on exponential functions, based on genetic programming, hybrid models, and finally, other proposals that do not fit well in the previous categories.

2.3.2.1 Gaussian Mixture Models

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements, such as vocal-tract related spectral features in a speaker recognition system. Thereby, voice quality estimation based on GMMs is an extended methodology explored by several authors.

Falk *et al.* [80]–[83] made an extensive use of GMMs. In their first works [80], [81], GMMs were used to generate artificial reference models of speech behavior. This technique compares distortion features introduced in these references with those affecting the real signal stream. Therefore, a double-ended quality estimation algorithm is emulated. In addition to these works, authors introduced an enhancement to improve the quality estimation accuracy when noise suppression algorithms are incorporated, by means of a fast approximation of the Kullback-Leibler Distance (KLB) between GM reference models of clean, noise corrupted, and noise suppressed speech and a GM model trained for the test speech signal [82]. Meanwhile, an additional work [83] proposed a modification to the GMM-based architecture described in aforementioned works. This modification consists on including additional information related to the transmission and coding schemes employed in the communication, showing a better performance in terms of accuracy. Results in discussed works showed better accuracy than P.563 in several scenarios, being also remarkable the decrease obtained in the quality estimation processing-time.

Following the same trend, Wang *et al.* proposed two different quality estimators. First, a quality estimation algorithm based on GMM and Support Vector Regression (SVR) was proposed [84]; then, an enhanced non-intrusive objective speech quality evaluation method based on Fuzzy Gaussian Mixture Model (FGMM) and Fuzzy Neural Network (FNN) was also presented [85]. In the former, authors used GMM to form an artificial reference model of the behavior of Perceptual Linear Predictive (PLP)

features of clean speech. Consistency measures between the degraded speech and the reference model were employed as indicators of speech quality. The effective least square SVR arithmetic was used to map the consistency values to the predicted MOS. In the latter work [85], an improved version of the previous method was proposed based on FGMM and FNN. FGMM was employed, instead of GMM, to form the artificial reference model. FNN regression algorithm was used to map the consistency values to the predicted MOS. The obtained results in both works outperformed the standard P.563 employing several coding schemes, such as G.711 and G.729; however additional tests under different conditions and multilingual databases would be necessary, in order to lead to more robust and comprehensive algorithms.

2.3.2.2 Neural Networks

It is well-known that Neural Networks (NN) have been extensively employed to emulate human behavior. Following this direction, several authors have developed quality-estimation models for voice communications based on Artificial Neural Networks (ANN) [86], Random Neural Networks (RNN) [87]–[90], and Self-Organizing Maps (SOM) [91].

Sun and Ifeachor [86] analyzed the effect on the call quality of four different parameters, namely, codec type, gender, loss pattern, and loss burstiness. In order to model the relationships between these elements and perceived speech quality, a neural network model was developed to learn the non-linear mapping from these parameters to a MOS score. Results showed good accuracy and correlation with PESQ, demonstrating that packet loss has a severe impact on perceived quality, and that female voices tend to be worse evaluated (more sensitive) than those of male talkers for the same network impairments.

In turn, Mohammed *et al.* made a deep study of their Pseudo-Subjective Quality Assessment (PSQA) technique [87], [88]. This model uses a RNN-based quality assessment mechanism in order to evaluate the impact of certain quality-affecting parameters, such as coding scheme, redundancy, packet loss rate, Mean loss Burst Size (MBS), and packetization interval, on real-time listening quality. PSQA allows evaluating the effects of these parameter interactions as a whole. The methodology employed in this model is as follows. First, a subjective MOS test is carried out, selecting a set of quality-affecting parameters. From this step, a match between the

different combination of impairment values and the quality experienced by the users is obtained. These data is used to train and validate a RNN; thus, PSQA works by learning how humans react to the communication from the quality point of view, through a set of selected variables. An enhancement of this model was also presented [89]. In this work the effect of delay, jitter, and Forward Error Correction (FEC) mechanisms were also taken into account. Consequently, PSQA was extended to assess not only the listening quality, but the conversational quality. In their results, authors showed that the key parameters affecting the conversational quality are packet loss rate, coding bit-rate, and the FEC mechanism. The impact of other impairments, such as delay or jitter, is low and subordinated to that of the packet loss process. In contrast to previously-discussed works, authors claimed that the mean loss burst size does not play a significant effect on the QoE. This fact was explained because of the good performance of the employed FEC mechanism and the PLC algorithm implemented in the codec used in their experiments. PSQA quality estimation showed good correlation with subjective tests.

Cherif *et al.* also made use of RNN to capture the non-linear relation between network parameters which cause voice distortion and the perceived quality [90]. The proposed model, so-called A_PSQA, receives as input just two parameters, namely, the packet loss rate and the Mean Loss Burst Size (MLBS) of the VoIP communication. The latter is employed, as in other reviewed models, to have an idea about the burstiness of losses. In this work, the effect of jitter is ignored as, with the absence of a de-jittering buffer, it is considered like additional packet losses. In order to train the RNN, authors developed a database of MOS scores for different speech samples transmitted under different loss conditions, characterized by the Gilbert model. These scores were attained by using the PESQ model. The employed coding schemes were iLBC and SPEEX codecs, with and without PLC. After training the RNN, results showed very good correlation with PESQ, beating the standard E-model and the IQX model, which will be analyzed in section 2.3.2.3. As stated in their conclusions, an extension of this algorithm including additional coding schemes would be desirable, as well as the inclusion of the delay impact on the quality of the VoIP calls.

Madhi and Picovici [91] presented an algorithm based on data mining (SOM) called Perceptual Output-based Speech Quality Evaluation (POSQE). POSQE is modeled using the following methodology. The degraded signal under study passes through a pre-process step in which it is segmented and each obtained frame is classified as voiced

or unvoiced. Then, only the voiced frames are analyzed and transformed into a speaker-independent perception-based parametric vector. Meanwhile, an artificial perception-based parameter is developed by using an internal reference codebook formulated from a number of clean speech records covering a wide range of human speech variations. Both vectors are correlated and the objective-auditory distance measured between the two vectors is computed. These operations are performed by a SOM. Finally, an appropriate logistic function is used to map the distance estimated into the corresponding objective listening MOS. Results show that POSQE outperforms the ITU-T PESQ in a large number of test cases and provides similar overall performance to that of the ITU-T P.563, reducing its computational complexity.

2.3.2.3 Exponential Functions

The IQX hypothesis, developed by Hoßfeld *et al.*, assumes an exponential functional relationship between QoE and QoS. Authors claimed that the more sensitive the subjective sensibility of QoE is, the higher the experienced quality. Under this assumption, they assumed that the change of QoE depends on the current level of QoE given the same amount of change of the QoS value. Thus, the IQX hypothesis is formulated as shown in (2.7),

$$QoE = \alpha \cdot e^{-\beta \cdot QoS} + \gamma \quad (2.7)$$

where α , β , and γ take different values depending on the codec employed, and the QoS parameter is measured in terms of packet loss rate, delay, and jitter.

In a first study [92], authors focused on the packet loss probability to measure the quality of service using a well-known commercial VoIP application; an extension of this work was then presented in [93], in which delay and jitter were also taken into account to assess the QoS. The results included in these works showed good accuracy, verifying the exponential relationship between QoE and QoS, but authors did not evaluate the combined effect of the different impairments assessed, e.g., packet loss rate and jitter, which could introduce important variations in the measured QoE. In addition, authors did not estimate the effect of bursty losses that, as demonstrated by other authors [45], [52], has a notable impact on the VoIP QoE.

Aiming to simplify the parametric approach as much as possible and at the same time further explore the burstiness-relating metrics as recommended in [94], Jung and

Mazano [95] presented three parametric models obtained through regression analysis, namely Model A, Model B, and Model C, whose unique input parameter is packet loss and the output a MOS value in the range 1-5. The simplest model, Model B, is expressed as shown in (2.8), where P_{loss} corresponds to packet loss probability.

$$Model_B(P_{loss}) = 3.108 \cdot e^{-0.06561 \cdot P_{loss}} \quad (2.8)$$

In turn, Models A and C are enhanced versions of Model B and include the effect of two additional burst-related metrics, specifically burst density and the fraction of burst loss within loss. Model C was designed as a quality model to be incorporated as VoIP QoE monitoring method in mobile energy-constrained end-points. For additional information regarding Models A and C, please refer to [95]. The accuracy of the models was validated using PESQ, but no comparison with IQX or any other parametric models was carried out.

2.3.2.4 Genetic Programming

Raja *et al.* [96], [97] presented a Genetic Programming (GP) based symbolic regression approach to derive a speech quality estimation model. The main advantages of GP, as claimed by the authors, are its capability of ignoring irrelevant network parameters while focusing on the most salient ones, and its production of human-readable results in the form of analytical expressions. Hereby, GP was employed to generate a mapping between the network and encoding parameters, and the quality perceived. The performance analysis of the GP proposal showed a good accuracy, outperforming the standard P.563 and obtaining good correlation with PESQ. However, additional evaluation with different network conditions and encoding schemes would be recommendable to ensure the generalization and robustness of the algorithm.

2.3.2.5 Hybrids Models

The approaches discussed above can be classified as signal-based models, parametric models, or packet-layer models. The former, estimate the QoE by processing directly the human speech, analyzing the distortion introduced in the voice signal. These models have shown to be sensitive to bursty packet loss and PLC algorithms. In turn, parametric models base its QoE estimation on the assessment of different impairments introduced by network and encoding schemes, which make them sensitive to background noises or noise suppression strategies. Finally, the packet-layer models

have not been fully developed, because of the difficulty of representing the complex interactions among the different impairments affecting the VoIP QoE just from parameters extracted from the packet headers. Thus, there is a current trend in which the best characteristics of all approaches are employed. These methodologies, so-called hybrid models, are gaining momentum and several proposals have been developed by different authors.

Jelassi *et al.* [98] extended the conventional parametric speech quality estimation models by considering the voicing feature of lost packets. This model builds a voicing-aware speech quality model that allows to accurately quantify the effect of lost packets according to their voicing property. By using multiple regression analysis, the speech quality estimation is obtained as a function of the different values of quality obtained for voiced and unvoiced frames, weighted by fitting coefficients. Further information is gathered from the packet loss pattern, distinguishing drops of voice and unvoiced frames. To do that, a sender-based notification scheme is adopted, in which additional data is introduced into the VoIP packets by the transmitter. This information allows the receiver to differentiate between voiced or unvoiced lost packets, but with an extra-consumption of bandwidth.

A listening-only model based on both network impairments (packet loss), and voice distortions (temporal clipping and noise) was presented by Ding *et al.* [99]. Authors developed a three-step algorithm that works as follows. First, the aforementioned impairments are detected. In order to evaluate (i) the occurrences of packet loss, (ii) the loss pattern and (iii) the employed coding scheme, an analysis of the IP headers is conducted. Additionally, a differentiation of silenced, voiced, or unvoiced packets is done, through voice payload analysis. The detections of temporal clipping and noise are also based on processing the voice payload. In the second step, the impact of individual impairments is quantified. Finally, the overall quality evaluation model is built by integrating individual impairments and employing the E-model. Results showed high correlation with PESQ (mapped into listening-only MOS), but several important impairments were ignored, such as delay, jitter, echo, etc. Additional work would be recommendable to extend this model to a conversational quality estimator, improving also its predictions by comparing them to those obtained by subjective tests.

In another work [100], Falk and Chan extended the use of their GMM proposals, discussed in previous sections, by integrating packet header analysis. GMM are used,

again, to generate an artificial reference model that is compared with the transmitted speech signal. On the other hand, a parametrical analysis similar to previous [98] is performed, evaluating the VoIP header. Consequently, this model inherits the limitations discussed above, related to the relevance of ignored impairments such as delay or jitter. A key characteristic for this methodology is its low computational complexity, which is 88% lower than the ITU-T standard P.563. Additionally, the proposed algorithm improves both, pure parametric approaches by measuring distortions that are not captured by connection parameters and pure signal based models by reaching lower per-call estimation errors. Thus, this methodology presents promising characteristics that could make it to be a solid alternative to the standard models.

2.3.2.6 Other Methodologies

Numerous works employing different techniques than those mentioned so far can also be found on the related literature. In this subsection, it is reviewed other quality estimation models which cannot be categorized in the previous sections but have served as the starting point to many other works.

One of the first single-ended models was presented by Jin and Kubichek [101]. This methodology makes use of a PLP-based Vector Quantization (VQ) technique to measure objective output-based speech quality. This algorithm works comparing the perceptual vectors extracted from the distorted speech-signal with a set of perceptual vectors composed by degradation parameters extracted from a dataset of clean reference signals. This speech reference codebook is built by the K-means algorithm and VQ-based distances are used to estimate speech distortion. In this model, three different distance measures are employed, namely, median minimum distance, transition probability distance measure and a combination of both. Authors tested their model under a variety of conditions, showing its effectiveness and robustness against speaker and content variations. However, the testing dataset employed was not as wide as desirable, in terms of accounted distortions. Consequently, the generalization of the algorithm was not deeply evaluated. The importance of this work lies in that it was the first single-ended quality estimator proposed; later, many other works, as those based on GMM [80]–[85], employed the same concept of generating an artificial reference model of speech behavior to compare it with the distorted received signal. Li and Kubichek also made use of this methodology but using a continuous hidden Markov model to replace the missing source speech record [102]. Euclidean and symmetric distances

were employed to measure the difference between the received and the reference speech signals. Authors obtained good correlation between the objective and subjective scores but, as in the previous work, only two datasets were used to verify the accuracy of the algorithm. Another approach based on hidden Markov model was also proposed by Talwar and Kubichek [103].

Regarding the different burst patterns affecting the packet loss rate of the VoIP stream, Roychoudhuri and Al-Shaer proposed an algorithm so-called Genome [104]. This methodology monitors at real-time the burst loss size and its preceding inter-loss gap in order to evaluate the different impact on the QoE of different loss distributions. Authors evaluated Genome through an extensive set of experiments, considering five different codecs and multiple loss patterns. They compared their results with those obtained by PESQ, obtaining accuracy among 96% - 98% in average. Finally, this model was also integrated in an adaptive multi-codec audio control mechanism, in order to achieve high levels of QoE during the course of the call, by performing coding-rate control.

With the aim of obtaining a light algorithm in terms of computational complexity, Jung *et al.* developed a QoE estimator just based on the measurement of the burst periods occurred during the call [105]. Authors employed the Gilbert-Elliot model to generate vectors including different loss conditions. In order to estimate the QoE of each particular vector, it was employed the PESQ model, obtaining a MOS characterization for each loss pattern by testing different-languages speech samples. By means of this analysis, authors obtained a series of curves related to each loss pattern employed. These curves were assumed as the quality estimation for any speech signal received with this particular packet loss pattern. Authors also provide a function for error estimation, depending on the burst characteristics. This model ignores a number of key parameters affecting the QoE of VoIP communications, aiming the simplicity of the algorithm and allowing embedding it in low capacity end-devices such as VoIP phones. For that reason, the model accuracy is quite low, and the generalization of the algorithm should be demonstrated with additional tests, employing different voice sources and loss patterns.

Following the same idea of obtaining a very light model which could be employed in end-devices, Mkwawa *et al.* just employed the Received Signal Strength Indicator (RSSI) to obtain an estimation of the transmission QoE. In a short work [106], authors

presented a mapping between the RSSI and the QoE on a Wi-Fi environment with a singular coding scheme, obtaining the relationship between RSSI and packet loss rate. The obtained mapping was barely accurate, with just 3 RSSI-to-QoE mapping-values which can be related with the tables provided in the ITU-T Rec. G.1010. In these tables, several mappings of packet loss probability onto QoE for different coding schemes are provided. Additional work would be desirable, extending the presented work for evaluating different coding schemes, or showing the computational load while the algorithm is working, as it is the unique advantage provided by this model in comparison with other proposals.

Another original proposal was described by Hassan *et al.* [107]. In this work, authors modeled QoE by measuring human irritation which, when reaches high levels, leads the communicator to quit the call. Authors presented three major contributions, namely, (i) employing the human effort as user QoE modeler to predict the (frustrated) user state that may lead to call quitting, (ii) formulating the provider's dilemma of allocating additional resources or not, as a non-cooperative game between the provider and the VoIP user experiencing deteriorating QoE, and (iii) taking into account the case of a "sophisticated" user, who may try to receive better service by generating fake efforts and discontent behavior. In their experiments, authors conducted a series of real calls, in which "human efforts", e.g., the total number of times the user exerted human effort, the timing of these efforts within the session, and whether the session ended naturally or prematurely, were assessed and compared with a final opinion score provided by the communicants. They obtained a high negatively correlation between the amount of human efforts during the call and its final QoE perception, but in a very small-scale experiments with just 58 calls evaluated. As a non-automatic methodology, this model is currently far to be actually implemented, as a human-supervisor is needed for each on-going call; for that reason, an extension of this work, showing some kind of automation on the human efforts recognition, would suppose a great step in the QoE estimation based on direct human-behavior observation.

To recapitulate, in this section a detailed review of the current standards and non-standard proposals related with objective non-intrusive VoIP QoE estimation has been presented. Focusing on the former, and considering the number of implementations and enhancements developed by the research community, the E-model has become the most popular methodology among the standards. Its ease of deployment and

comprehensibility are the most valued features by the research community. Regarding non-standard methodologies, many of them have been developed with the aim of improving the accuracy of the standards. Most of the reviewed alternative models have been compared with standard models, beating them on accuracy and/or computational efficiency. Although there is not any work comparing the characteristics of different alternative models, the final decision of employing one model or another clearly depends on the particular characteristics of the system under evaluation. Finally, the focus was on the hybrid models, which make use of a combination of diverse approaches. These models employ great number of parameters from different network layers; consequently, in some cases, cross-layer techniques are needed to gather the different metrics taken from several sources. The complexity of joining the impact of the different impairments affecting the communication is a challenge to beat and could suppose a significant improvement in the QoE of the VoIP transmissions. Hence, the paradigm of gathering different-layer metrics is currently gaining momentum due to its high levels of accuracy, and this seems to be the future direction to follow for the definition of a full compact QoE model according to the characteristics of modern VoIP services

2.4 IPTV QoE Estimation Methodologies

As shown in previous sections, for simpler communication services such as VoIP, QoE assessment is at advanced stage. However, it has been difficult for researchers to find out a QoE evaluation model that accommodates to more complete multimedia services due to their inherent complexity. One of these multimedia services with a notable market growth is IPTV/Internet TV [108]–[113]. IPTV is defined by the ITU-T IPTV Focus Group as “multimedia services such as television/video/audio/text/graphics/data delivered over IP based networks managed to provide the required level of QoS/QoE, security, interactivity, and reliability”. IPTV hence differentiates from cable or satellite TV because of the provision of full interactivity, high personalization, and flexibility. On the other hand, some researchers identify the difference between IPTV and Internet TV as that in IPTV, packets are not transmitted over the public Internet but travel solely over carrier's own servers and network. Nevertheless, two facts are making these two services to merge. First, the binding IPTV-TV set and Internet TV-PC has turned limited with the current emergence of smart TVs, which allow surfing the web and accessing to several on-line services and

apps based on video delivery from the Internet. Second, gadgets as tablets, smartphones, or laptops, are daily-used to consume all-kind of multimedia services, bringing new experiences to the user and being possible to find apps in the market for these gadgets that allow accessing to IPTV services using different consumer electronic devices other than the traditional TV set. As a consequence, although both technologies are in their infancy from a market-deployment perspective, both are heading into the same direction. It is in this context where a strong effort is being made to find a QoE assessment method for IPTV/Internet TV services that could ensure that users will be satisfied with end-to-end service performance.

Current proposals follow two different trends: (i) methodologies that estimate the quality of the multimedia service only focusing on the audiovisual experience, which depending on the source of information used to obtain their quality estimations can be classified into packet-layer models, bitstream-level models, media-layer models, and hybrid models; and (ii) proposals that evaluate the service quality from a wider perspective, i.e., in addition to the audiovisual perception they also take into account other key parameters that could impact user experience such as usability, zapping time, security, reliability/availability, or acceptability, among other factors. In the following sections, a comprehensive survey on QoE assessment methodologies for future TV services including both standard and non-standard proposals is presented.

2.4.1 Standard Models

In this section, the efforts made by the ITU standardization body aiming to evaluate QoE in IPTV are described. A complete review of QoE efforts from the ITU was presented in [114], [115]. However, additional work has been done since the date of publication of these works that it will be hence included here.

The ITU-T has recently developed two new assessment methods for multimedia transmission, specifically for audiovisual streaming [116], [117]. Additionally, different Recommendations setting user requirements and multimedia-services performance objectives were developed [16], [18], and currently are employed as guides to develop a standard IPTV QoE estimator. Another Recommendation that provides key considerations about the end-to-end performance estimation for data applications is ITU-T Rec. G.1030 [118]. As stated in this Recommendation, the users' perceived performance of data applications on packet networks is dependent on many factors,

some key ones: end-to-end performance of packet network, application's dependency on the communication network, performance of the supporting devices, and user's task (user interaction with the application). These factors are taken into account to guarantee user satisfaction. A model of data application performance should include as many of these factors as possible, so once the application performance has been estimated, perceptual models can be applied to interpret the level of end-to-end performance attained. In order to estimate the end-to-end performance in IP networks for data applications, first it is needed a network performance assessment in terms of packet transfer performance parameters, e.g., as defined in Rec. Y.1540 [17] or in IETF standards. Then, an application performance assessment should be obtained, followed by the perceptual performance assessment and the end-to-end performance estimation. ITU-T Rec. G.1030 [118] provides some perceptual models (web and http file transfer with TCP) to obtain the desired end-to-end performance assessment, but no further models are proposed for IPTV or multimedia services.

Focusing on the ITU-T proposals, a notable effort is being done by the ITU-T Study Group (SG) 12 "Performance, QoS, and QoE", the SG13 "Future networks including cloud computing, mobile and next-generation networks", and the SG16 "Multimedia coding, systems and applications" in order to settle the basis for QoE evaluation of IPTV services. Whereas SG13 and SG16 are study groups with QoS related technical activities, SG12 is the core group that focuses on standards that accomplish end to end QoS/QoE provisioning in IP environments for a wide variety of networks and applications.

To start with, three Recommendations specifically designed for the IPTV service are highlighted: ITU-T Rec. G.1080 "Quality of experience requirements for IPTV services" [119], ITU-T Rec. G.1081 "Performance monitoring points for IPTV" [120], and ITU-T Rec. G.1082 "Measurement-based methods for improving the robustness of IPTV performance [121].

G.1080 defines QoE requirements from an end user perspective (specified as end to end) and determines how they influence network transport and application layer behavior. Specifically, G.1080 considers the following issues in the process of determining QoE performance targets: purpose of the IPTV service, QoE level of the current broadcasting system (which sets user expectation), compression coding scheme to be used for the service, content characteristics, content provider requirements, and

customer satisfaction. Having this in mind, there are a number of system performance characteristics that contribute to the QoE of the media stream. For instance, the codec and the encoding bit-rate, media resolution in the source and at the display, corruption or loss of information, delay, or video content; all of them contribute to a high variability in the perceived quality of the video output. Additional factors that influence viewers' response are cultural background, motivation, emotional state, previous experiences, how much the user is paying for the service, what special benefits (if any) it provides, and so on. According to G.1080, direct evaluations of QoE are designed to exclude these last factors, since they are not generally under the control of a network operator, and so do not contribute to equipment requirements. But even though they do not impact on equipment, it is questionable whether they should be taken into account or not. An interesting observation is that acceptability is not equivalent to QoE, i.e., a good level of QoE could be perceived, but not being acceptable because of the price or because of expectations [12]. Although both terms are clearly not equivalent, it is undoubtedly worthy to take acceptability into account when evaluating QoE, since it is not useful to have a service with a high QoE if no user is going to pay for it. With this in mind, ITU-T Rec. G.1080 addresses provisional QoE targets and shows how to express QoE requirements only in the context of numerical parameters such as bit-rate or packet loss rate.

Regarding QoE requirements for media compression and synchronization in IPTV, the main factors identified by ITU-T Rec. G.1080 influencing video QoE at the application layer due to compression are: quality of source materials, baseline quality of the codec, resolution, bit rate, application layer video encoding (Constant Bit Rate versus Variable Bite Rate at the encoder output), group of pictures structure, motion vector search range, rate control, pre-processing, and tandem encoding and rate shaping (e.g., digital turnaround). Because of the importance of audio/video synchronization, tighter tolerances than typically used for video conferencing applications are being recommended. Particularly useful are the recommendations provided by G.1080 on minimum bit-rates for several video and audio codecs at the application layer (prior to IP encapsulation), specifically for standard definition TV (SDTV), video on demand and premium content within SDTV, and high definition TV (HDTV). As an example, Table 2.5 shows recommendations for HDTV. The impact of network transmission performance on QoE is given by loss, latency, and jitter. In general, reasonable end-to-

TABLE 2.5. Minimum levels of transport layer performance required to provide satisfactory QoE for H.264, SMPTE 421M, or AVS encoded HDTV services (extracted from [119])

Transport stream bit rate (Mbps)	Jitter (ms)	Maximum duration of a single error event (ms)	Corresponding loss period in IP packets	Loss distance (events per 4 hours)	Corresponding average IP video stream PLR
8	<50	<=16	<=14	<=1	<= 1.28E-06
10	<50	<=16	<=17	<=1	<= 1.24E-06
12	<50	<=16	<=20	<=1	<= 1.22E-06

end delay and jitter values are not problematic due to set top box de-jitter buffers, provided that the de-jitter buffer size is set to match network and video element performance. Video streams however are highly sensitive to information loss, and in turn, the QoE impact is correlated to a number of variables including: type of data lost, codec used, MPEG transport stream packetization used, loss distance and loss profile, and encoding bit rate. Special attention is needed in the case of Smart TVs or mobile devices, situations not covered in G.1080. Buffers optimization for future TV services in these elements are under study [122], [123].

In IPTV, QoE is also influenced by channel zapping and tricking modes (e.g., pause, play, rewind, etc.), which are called control functions. Channel zapping time has a strong relationship with QoE. It is mainly determined by the time required to have a proper frame at the set top box to start decode processing for the new channel. Similarly to web navigation, it would represent the time elapsed between going from one site to another. As a QoE parameter, channel zapping time can be described by three components: IGMP (Internet Group Management Protocol) delay, buffering delay, and decoding delay. The explicit relation between the channel zapping time and the user perceived quality expressed as a MOS is still under study as we will discuss in the following sections. In turn, Video on Demand (VoD) trick modes provide VCR-like features in VoD services. Each mode has its own delay, which has to be low enough to meet QoE user requirements, and whose impact on QoE also needs further study [124], [125]. Metadata (electronic program guides, caption, parental control, etc.), browser, and content navigation are the last IPTV services taken into account in the QoE analysis proposed by G.1080. Metadata main implications are availability and correctness; for instance, having always proper movies age-rating. For the other two services, ease of use is the main parameter with a clear effect on QoE.

Secondly, once the parameters that affect QoE in IPTV services are identified in ITU-T Rec. G.1080 (rate and type of data lost, coding scheme, packetization time-length, loss pattern, delay, and jitter, etc.), ITU-T Rec. G.1081 [120] defines performance monitoring for IPTV. IPTV performance monitoring can be software based, hardware based, or a hybrid combination. Fig. 2.8 shows a general topology divided into specific monitoring domains. As recommended by G.1081, monitoring should be done at each domain boundary, so that a complete end-to-end monitoring view can be obtained. The elements identified in Fig. 2.8 are the content provider, the service provider, the network provider, and the end user. It is also recommended the use of a performance monitoring management platform in order to interpret collected data from the IPTV system. Interfaces among the platform and the monitoring points are still under development. The benefits of doing performance monitoring as enumerated in G.1081 are: error detection in an end to end system, checking resource utilization and work load of system components, comparing metrics regarding performance of different system deployments, getting a starting point for system modeling, bottlenecks discovery, IPTV network deployment optimization, and ensuring no degradation of system performance with time.

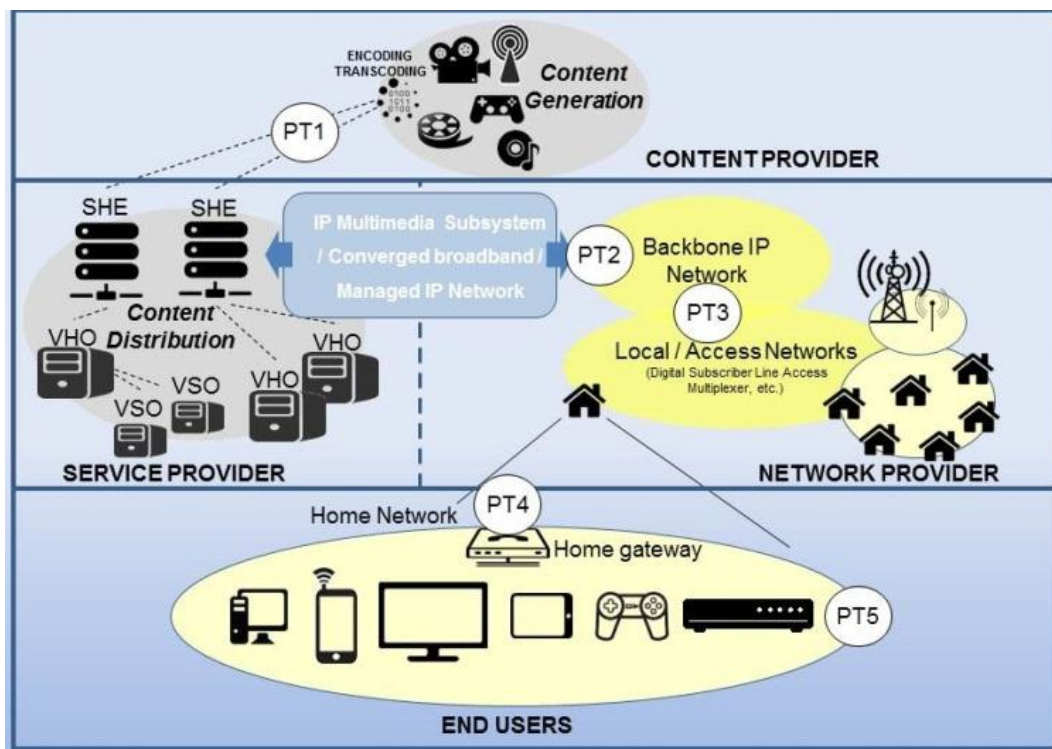


Figure 2.8. IPTV performance monitoring points as recommended in G.1081 [120].

TABLE 2.6. Information capture at each monitoring point (extracted from [120])

Monitoring point	Information captured and reported after being processed
PT1	<ul style="list-style-type: none"> • Coding and quality parameters of video, audio, and text content • Acknowledgement for the report of source content transfer receipt
PT2	<ul style="list-style-type: none"> • Identifier of media flow • General program attribute (service type, total bit rate) • Video parameters (bit rate, encoding format of video content, resolution of video content, GOP parameters, timestamp of I frames) • Audio parameters (number of audio channels, bit rate and encoding format of audio channels) • Video, audio, and multimedia quality (MOS) before sent to the network*
PT3	<ul style="list-style-type: none"> • Identifier of media flow • IP network capacity information(Y.1540, RFC5136) • IP network performance (Y.1540, Y.1543, Y.1544). Mean one-way delay, one-way packet delay variation, packet loss ratio, packet loss profile, path unavailability. • IP multicast network performance (Y.1544). Successful join time, successful leave time, group mean one-way delay, group IP service availability, mean group loss ratio. • One-way loss pattern related performance (Y.1540, RFC 3357) • Video, audio, and multimedia quality (MOS) in the network domain*
PT4	<ul style="list-style-type: none"> • Same as PT3 but from the edge point of IPTV (e.g. home gateway or set-top box) • Access to raw IP information of the audiovisual or text stream
PT5	<ul style="list-style-type: none"> • Video, audio, and multimedia quality (MOS) perceived by end user*

Depending on the location of the monitoring point, different parameters will be monitored (please see Table 2.6). PT1 checks the source audiovisual quality and metadata verification. PT2 monitors original streaming quality. PT3 is in charge of IP related performance parameters. PT4, located for instance at the set top boxes, focuses on quality of streaming, audiovisual quality, and IPTV service attributes. Finally, PT5 directly relates to end-user QoE, and monitors audiovisual quality, text accuracy, and IPTV service attributes from the end-user perception. Notice that ITU-T Rec. G.1081 only defines monitoring methods for PT3 and PT4 because they are concern with IP network performance, which was defined in ITU-T Rec. Y.1543 [126] and ITU-T Rec. Y.1544 [127]. The list of parameters to be measured or monitor at each point (as shown

in Table 2.6) is not closed. Other parameters such as terminal metrics (e.g., type and size of display, memory use, energy consumption, etc.) or user behavior (e.g., spent time, zapping frequency, etc.) could be of interest from the monitoring perspective [128], [129]. Also, the production of data available will undoubtedly increase with the advent of Internet of Things, and how to effectively manage, mine, and use this data to improve QoE monitoring and analysis is decisive [130], [131].

Thirdly, ITU-T Rec. G.1082 called “Measurement-based methods for improving the robustness of IPTV performance” aims to improve the IPTV service from a QoE perspective and to provide methods to do so based on real-time measurements (taken from the monitoring system explained in ITU-T Rec. G.1081). Fig. 2.9 illustrates the main concept. Due to the usually different nature of service and network providers, they could manage a different Monitoring, Management and Performance Analysis Platform (MMPAP), exchanging total or partially collected data, which requires collaboration between providers. This could suppose a key issue to deal with, due to the need of using common standards and methodologies. Based on the measurement data collected, the MMPAP infers QoE metrics. If these are below a predefined trigger, then adjustments procedures are carried out aiming the improvement of the QoE provided. There are several ways to proceed with the adjustments. In case of media resource control, the operations available could be adjusting the encoding bit-rate based on parameters like available path capacity or expected QoE, and encoded layer adjustment if the media flow has been encoded into multi-layers. Regarding the network resource control, the system could be able to modify resource allocation (e.g., degrading less critical

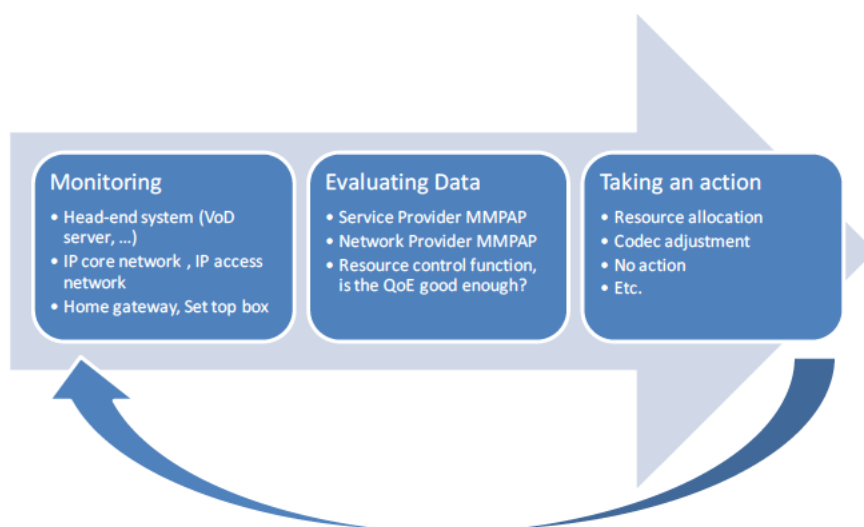


Figure 2.9. Simplified general process to improve IPTV QoE based on real-time measurements (MMPAP ≡ Monitoring, Management and Performance Analysis Platform).

services), to increase or decrease the Forward Error Correction (FEC) redundancy of the Real Time Protocol (RTP) packet depending on packet loss ratios, or to re-establish the delivery path (e.g., using Multi-Protocol Label Switching techniques).

At this point, the following questions rises: how to set up the appropriate triggers? As indicated previously, ITU-T Rec. G.1010 defines QoS requirements, establishing maximum tolerated limits for packet loss rate and delay in different multimedia applications, such as conversational audio/video, video streaming, or data applications, among others. However, this methodology does not take into account the inter-relations between impairments, which could cancel each other their effects. For that reason, the use of a single index (e.g., MOS) as trigger, gathering the impact of all impairments affecting the transmission would be an interesting solution. This approximation has been recently released in the form of a new set of ITU-T Recommendations, focused on parametric QoE monitoring of IPTV services [116], [117], [132]–[135].

ITU-T Rec. P.1201 [116] called “Parametric non-intrusive assessment of audiovisual media streaming quality” or P.NAMS (Non-intrusive parametric models for the Assessment of performance of Multimedia Streaming), describe video, audio and audiovisual quality models. ITU-T Rec. P.1202 [117] called “Parametric non-intrusive bitstream assessment of video media streaming quality” or P.NBAMS (Non-intrusive Bit-stream models for the Assessment of performance of Multimedia Streaming) provides only video quality estimations. The P.NAMS models just employ packet-header information, while the P.NBAMS models make use of further bit-stream information, such as coding-related information. Although P.NBAMS models can be more accurate than P.NAMS models in their quality predictions, they require more computational power to estimate the video quality given their higher complexity. Both P.NAMS and PNBAMS deliver their quality estimations in a MOS scale. Therefore, P.NAMS outputs three different MOS scores on the 1 to 5 scale: estimated audiovisual quality, estimated video-only quality, and estimated audio-only quality; and P.NBAMS just provides a MOS value on the 1 to 5 scale with the estimation of the perceived video quality.

These Recommendations [116], [117] have not been developed as general algorithms, which could be employed under any network and coding conditions; instead, the models have been developed for a set of dedicated service implementations, which are assumed to be meaningful representations of today's IP-based streaming

video services, such as mobile TV (low resolution application area [132], [134]) and IPTV (high resolution application [133], [135]). The PNBAMS models for high resolution are still under study and hence not completed [135]. It is interesting to note that there are some applications and test factors the models are not intended for, such as video streaming services with significant rate adaptation (e.g., dynamic adaptive streaming over HTTP, also known as DASH). Both Recommendations can work in four operation modes: static operation (NN), non-embedded dynamic operation (BN), non-embedded distributed operation (CN), and embedded operation (CC), although currently just the latter is supported in the current models implementations. In the first three modes, the model is located in the network, whereas it is located in the client in the last mode. Measurements can be done at the client, the network, or both, being indicated by the first letter of the acronym (N, network, C, client, and B, both). Each operational mode assumes different inputs and locations along the transmission chain as summarized in Table 2.7. Even though these methodologies are based in the information extracted from the packet-header or the bit-stream, additional inputs can be also received in order to obtain more accurate quality estimations. This extra-information is collected by the P.NAMS/P.NBAMS models from different interfaces, depending on its source. Both Recs. have three main inputs, two in common, which are buffering information and static media-and-decoder information, and one specific for

TABLE 2.7. P.1201 and P.1202 operational modes (extracted from [116], [117])

Class	Name	Description
Mid-point or End-point	NN	The model uses information from the local transport layer, prior knowledge about coding and end-point
Mid-point	BN	The model uses information from the local transport layer, prior knowledge about coding and information about the end-point collected through measurement reporting protocols
Mid-point	CN	The model, located inside the network, uses information from the transport layer measured at an end-point and collected through signalling protocols, prior knowledge about coding and information about the end-point collected through signalling protocols
End-point	CC	The model uses information from the local transport layer, information from the end-point, and prior knowledge about coding

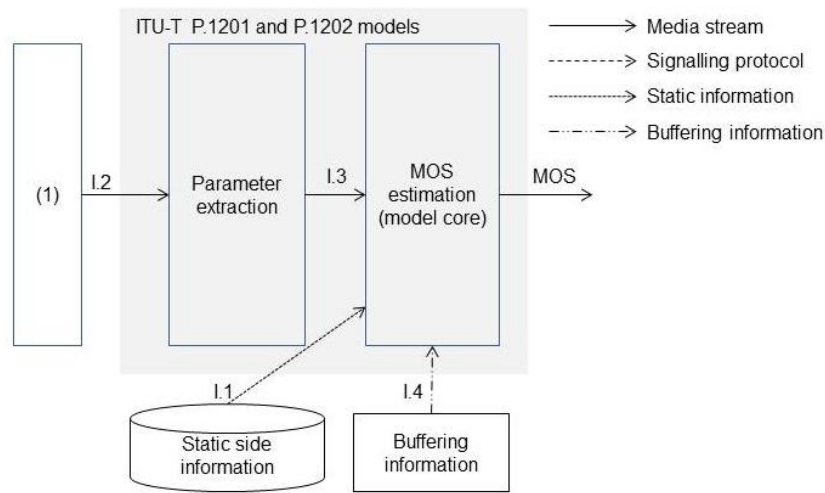


Figure 2.10. Overview of the ITU-T P.1201 and P.1202 model interfaces. (1) In case of P.1201 it is packet header information (dynamic), in case of P.1202 it is encoded bitstream. (I \equiv Interface) (extracted from [116], [117]).

each one, which is packet-header information for P.NAMS and encoded video-stream for P.NBAMS (see Fig. 2.10). Examples of static information about the media stream and the decoder are audio/video destination ports, video codecs, video resolution, video frame rate, video packet loss concealment, and so on, and it is available out-of-band from packet information or from a player API. Buffering information or rebuffering information/parameters (e.g., buffer length, buffer start time, etc.) can be obtained from the media buffer in the client side or estimated by a buffering estimation module using packet information. PCAP (Packet CAPture) files are expected to be used for both packet-header information and encoded video bitstream, being necessary to include the payload in the latter.

Fig. 2.11 shows the block diagram for the audiovisual estimators (P.1201) and Fig. 2.12 represents the block diagram employed by the only-video estimation platform (P.1202). Please, see the details of each block in their respective Recommendations. Notice that the output provided by both models is different: three quality estimations (audio, video, and audiovisual) are provided by the P.1201 models, while just one video quality score can be obtained from the P.1202 models. Note also that in the P.1202.1 model (Fig. 2.11), the input must be a video stream coded with ITU-T H.264 codec; additional extensions for this model would be necessary in order to generalize the use of this model to any other coding scheme.

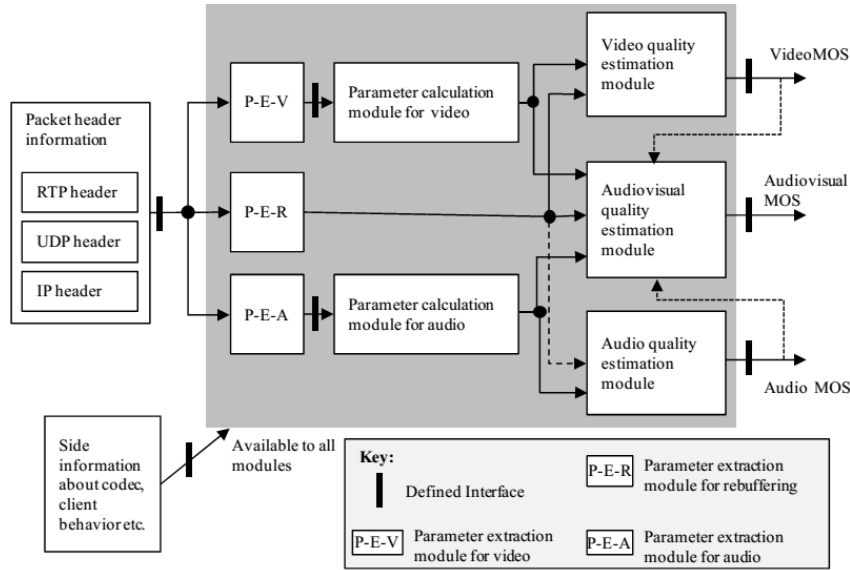


Figure 2.11. Block diagram for the ITU-T P.1201.1 and P.1201.2 models (extracted from [132]).

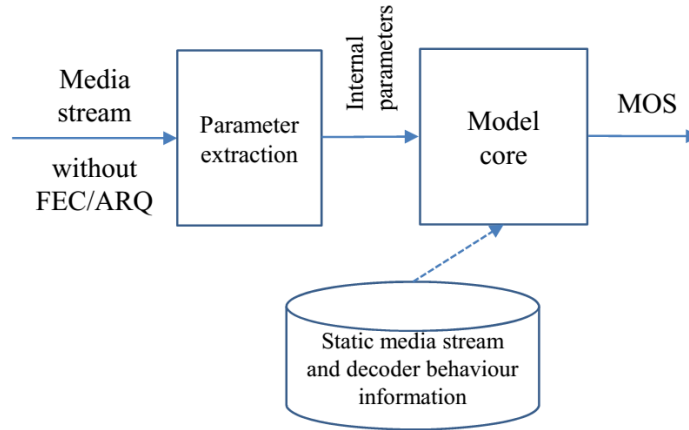


Figure 2.12. Block diagram for the ITU-T P.1202.1 and P.1202.2 models (extracted from [117]).

2.4.2 Non-Standard Models

IPTV QoE estimation methodologies can be grouped in two different families, depending on the factors taken into consideration to assess the service quality. First, there are a number of models that estimate the quality of the multimedia service focusing just in the audiovisual experience (section 2.4.2.1). According to the source of the information employed to obtain their quality estimations, these methodologies can be classified into the following categories: media-layer models, packet-layer models, bitstream-level models, and hybrid models. On the other hand, it can be found different proposals that evaluate the service quality from a wider perspective (section 2.4.2.2), i.e., in addition to the audiovisual perception, they also take into account other key

parameters which could impact user experience, e.g., zapping time, security, reliability/availability, acceptability, etc.

In the following, it is provided an overview of different alternative methodologies to the standards, which can be found in the related literature.

2.4.2.1 Audiovisual-only Quality Assessment Models

These models are focused on evaluating just the provided video/audio quality ignoring any other factor, which could alter the customer's service perception. As aforementioned, there are several approaches, depending on the information considered to obtain their estimations: media-layer models, packet-layer models, bitstream-level models, and hybrid models.

The **media-layer models** were the first attempt to simulate subjective tests. For that reason, a great number of different proposals can be found in the literature. These models decode the multimedia signal and analyze its quality degradation pixel-by-pixel. Thus, these approaches are usually based on a human vision model, providing good-level of accuracy because they evaluate the signal that the user actually receives. However, these techniques need additional information about the packet loss suffered by the media stream to locate the impaired parts in the distorted video. Depending on the information available in the measurement point, these models can be categorized as full, reduced, or no reference models. In the following it is reviewed those works with bigger significance, either by their correlation with subjective analysis or by the open road and proposals arising from them.

In [136], two different video quality estimation models were proposed. The first one is a full reference model based on the measurement of structural distortions. By means of statistical-feature extraction, the quality of each frame is evaluated. The overall frame quality estimation is obtained by averaging the quality assessed for the three color-components of the frame, namely, Y, Cr, and Cb. Then, the quality value for each frame is adjusted by introducing the effect of two factors: blockiness and motion. Finally, the average quality estimation for each frame is taken as the video-sequence quality value. The other proposed methodology was a non-reference model in which the video input is decoded in order to extract the inverse quantized Discrete Cosine Transform (DCT) coefficients, the quantization scale, and the motion vector for each block. Then, quantization errors of all the DCT coefficients are estimated and averaged together,

providing an estimate of the overall quantization error of the frame. As in the full-reference proposed model, the effect of blocking and motion is also included into each frame quality-estimation. Finally, the average of all these frame quality estimations gives the full sequence quality estimation.

A quite similar approach that also employs structural distortion as an estimation of the perceived visual distortion was presented by Wang *et al.* [137]. This full-reference model employs the Structural Similarity Index (SSIM) [138] to compare the structural differences between two images. This comparison is done accounting for the average and the covariance of the compared signals, which are considered rough estimates of the luminance and the contrast of the signals, respectively. The SSIM is successively applied to regions of 8x8 pixels and the average of these regions quality estimations are taken as the final image quality value. In order to apply this methodology to video sequences, authors divided the process in three levels. First, the SSIM is applied to the three colors composing a single frame (Y, Cb, and Cr). As mentioned above, this process is done in 8x8 regions of the image but now, aiming to reduce the computational load, just some regions are evaluated. The weighted sum of these values composes the single-frame quality estimation. Finally, the average of the single-frame estimates gives the final quality estimation of the video-sequence. As stated by the authors, the principal drawback of this algorithm is that it is focused on the comparison of still images. For that reason, motion features of video sequences are not considered. Additionally, as this model always employs average values, some important punctual impairment could be masked and not reflected in the final quality estimation; this could be a problem because this situation does not usually happen in subjective tests, in which important artifacts, although isolated, are normally considered.

Lee and Sim presented a model which analyzes the visual features of digital mobile videos focused on the edges and block boundary regions because, as they stated, those regions are more sensitive than others to the human visual system [139]. This full-reference model, so called KVQM (Kwangwoon Visual Quality Metric), extracts image features to evaluate the effect of a number of degradations, namely, edginess, blockiness, and blurriness. Three single values representing each of these features are calculated. The final quality estimation is obtained through the weighted sum of these three values. The comparison of the model performance with that of the EPSNR (Edge Peak Signal-to-Noise Ratio) model [140] shows promising results, but an extension of

this work regarding other important factors such as the content of the sequence or the motion features of the video would be recommendable.

Work in [141] presented a computational-friendly quality estimation methodology based on the correlation between the received video-sequence MOS and the Mean Squared Error (MSE) between the original and the degraded video signals. Authors proposed that MOS (in a 0-1 scale) decreases with the rise of the MSE weighted by the slope of the regression line, which is calculated automatically from the sequence edge strength. Similar to the model presented in [137], first the quality estimations are made separately for the different blocks composing the images. Then, these values are successively averaged to obtain the full frame quality estimation and, finally, the full video sequence quality value is obtained. Results showed better accuracy than PSNR [142], SSIM [138], and PSNRPlus [143] in terms of both correlation with subjective tests (Pearson Coefficient and Outlier Ratio) and computational time.

In turn, the **packet-layer models** only employ the information that is extracted from the different headers within the multimedia packet, e.g., IP/UDP/RTP. Thus, these models consider parameters such as packet loss, delay, coding scheme, etc. to obtain their quality estimations. These methodologies can detect the location of distortions in the multimedia stream and their impact on the audiovisual quality; however, they ignore important factors such as the type of multimedia content streamed, which has a significant impact on the user quality perception [144]. Due to the low computational requirements of these techniques, they are usually employed for in-service monitoring. This is because of the algorithms used to provide their estimations, which do not need to extract information from the media-signal through a heavy decoding process. Several examples of this approach can be found in the related literature [145]–[151].

Work in [145] presented a model that takes into account the distortions related to the encoding algorithm and packet losses. First, a parameter-extraction block is employed to analyze the packet header in order to obtain key information such as the codec employed, packet losses and the length of the packet-loss bursts. Then, a quality estimator focused only on the impairments introduced by the coding scheme is employed. This block estimates the quality of the video from the coding bit-rate by using a series of coefficients empirically calculated from each codec and bit-rate used. These coefficients were extracted from subjective tests conducted previously. Finally, the effect of packet loss is introduced in the previous quality estimation. Authors found

that no-matter the encoding bit-rate used, the relationship between the subjective-video quality and the packet loss ratio is always the same and can be modeled by an exponential function. Hence, the final expression to estimate the quality of the video signal in terms of MOS is that shown in equation (2.9),

$$V_q = 1 + \left(v_1 - \frac{v_1}{1 + \left(\frac{Br}{v_2} \right)^{v_3}} \right) e^{-\frac{PLR}{v_4}} \quad (2.9)$$

where v_1 , v_2 , v_3 , and v_4 are the aforementioned experimental factors that depend on the coding scheme applied. Results showed good accuracy with subjective tests with a cross-correlation value of 0.968 and RMSE (Root Mean Square Error) of 0.287. Because of the good levels of accuracy showed by this methodology, it was taken as the base for the video quality estimation model described in the ITU-T Rec. G.1070 [152] and it has been also employed in different works as video QoE estimator. However, some drawbacks are found in this model. First, due to the empirical nature of this model, a lot of work is necessary in order to obtain a coefficients data-base big enough to include the big amount of coding schemes and bit rates available. In [145], authors only obtained the coefficients for the H.264 codec, while in the Rec. G.1070 codec MPEG-2 and MPEG-4 were also considered. In addition, only two parameters are taken into account, namely, randomly distributed packet loss and coding bit-rate. It is well known that further factors could also impact IPTV QoE, e.g. end-device's screen-resolution, motion characteristics of the transmitted video-content, packet loss pattern, etc.; for that reason, an enhancement to this model can be found in [146]. In this work, the authors proposed some corrections to the expression (2.8) aiming at including the effect of other parameters such as the employed coding scheme (H.264 or MPEG-2), the video resolution, or the motion level of the video-content. Thus, the improved model presents more versatility than the original one, as it is capable of working with different kind of video sources. Another extension for Yamagishi and Hayashi's work [145] can be found in [147]. The presented methodology, so-called T-Model, includes the effect on the video quality of the packet loss pattern. To achieve that it was analyzed some characteristics of the packet loss bursts, namely, density, duration, and number of burst periods. In their results, authors showed improved accuracy under burst loss conditions than the original G.1070 model [145], [152].

Another approach to estimate the video QoE from the packet header fields is that presented in [148]. In this case, the model is focused on evaluating the video quality by the analysis of the error-concealment effectiveness and the errors location in each frame. In this work, authors evaluated the error-concealment on the basis of motion-level information. This was calculated from the B-pictures data-size. The smaller the data-size of these pictures, the lower level of motion, and, consequently it is considered a high effectiveness of the error concealment. The quality-score obtained from this stage is, then, adjusted considering the location of the error into the picture. As discussed in previous sections, an error in the center of the screen has greater impact on the perceived image quality than that located in the frame edges. The location of the video-quality degradation is estimated by the distance between the bit-position of a transmission error and the bit-position of picture start or end. As in [145], a subjective test was conducted in order to extract some empirical values, which were thereafter employed in the quality estimation calculations. Although results showed a high correlation coefficient of 0.853 with subjective measurements, the tests conducted were very limited, with just one studied codec (H.264), very few encoding bit-rates and only 5-seconds test-videos under consideration. In addition, just one error concealment algorithm was used, so additional test would be necessary to check the model performance under different conditions.

An algorithm meeting the requirements of the ITU-T Rec. P.1201 was presented in [151]. This model takes under consideration the following quality degradations: (i) compression artifacts due to encoding process, (ii) packet loss artifacts (including slicing and freezing artifacts), and (iii) re-buffering artifacts happened in the receiver side. The proposed algorithm first extracts from the packet header different parameters regarding video and audio quality. These parameters are passed to the next stage in which they are processed and prepared for the quality estimation modules. In this last stage, three different quality estimation modules provide audio, video, and audiovisual quality values in terms of MOS. Regarding audio quality estimation, the bit rate and lost audio frame length are used. In addition to these degradations, the impact of re-buffering is also taken into account for the video quality estimation. Furthermore, some experimental parameters are also employed in both calculations. Finally, based on these audio-only and video-only quality values the audiovisual quality estimation is derived. The presented results showed high correlation with subjective tests; it was attained

Pearson coefficients of 0.852, 0.830, and 0.941, for audiovisual, video, and audio quality estimations, respectively. For that reason, this model was selected as the standard algorithm included in the ITU-T Rec. P.1201.1 (P.NAMS-Low Resolution) [132].

Also following the guidelines of the ITU-T Rec. P.1201 [116], Raake *et al.* proposed the T-V-model [149]. This model classifies the input parameters in three categories. The “send” parameters take into account encoding information, e.g., codec employed, video format, etc. The “transport” inputs include network parameters affecting the video quality, such as packet loss, throughput, etc. Finally, the “receive” parameters gather the impairments caused by the receiver equipment, e.g., buffer behavior, error concealment, decoding process, etc. As stated in Rec. P.1201, the T-V Model provides four different quality estimations: audio, video, multimedia qualities, and a model of the interaction between audio and video quality. The calculations employed to obtain the video quality are similar to those employed in the E-Model [33]; that is, an additive scale in which several impairments caused by different sources decrease the maximum video quality, i.e., the quality of the source video before its insertion in the transmission system. These impairments are related to the errors caused by (i) transmission (e.g., packet loss), (ii) video-coding process, (iii) picture resolution (in case of resolution changes), and (iv) display and processing steps (e.g., de-interlacing). In this work, the two latter impairments are set to a value of zero because of the lack of a model describing them accurately. The presented results showed high accuracy in terms of correlation with subjective tests. Extending the findings in [149], the audio-only and the audiovisual modules are further described in [150]; again, the concept of impairment factor approach is employed for the audio module. In this case the predicted audio quality is based on the impact of audio compression and audio frame loss. In turn, a weighted combination of video and audio modules outputs is used to obtain the audiovisual quality estimation. Thus, the complete model is presented and validated using a large database of subjective tests. Due to high performance results, both in terms of Pearson’s Correlation coefficient and Root-Mean-Square-Error, compared to subjective tests, this model was finally selected as the winner of the P-NAMS-High Resolution competition and it was included as standard algorithm in the ITU-T Rec. P.1201.2 [133]. Deeper details of this model can be found in this Recommendation.

In turn, the **bitstream-level models** receive as input the same parameters as the packet-layer models in addition to the compressed multimedia stream. Thus, these methods have access to the video content feature and the detailed encoding parameters by analyzing the video bitstream. These models are more complex than the previous category, so they are not suitable to be installed in end-user devices. Depending on the analyzed parameters, several methodologies have been developed [153]–[156].

In [153], authors presented an active monitoring system that measures perceived quality by means of the previously defined Pseudo Subjective Quality Assessment (PSQA) technique [157]. PSQA is a model based on Random Neural Networks (RNN). Prior to the quality evaluation stage, the RNN undergoes a training process in order to capture the relationship between the parameters that cause the distortion and the actual perceived quality. This work focused on just two parameters affecting the video quality: loss rate of video frames and mean size of loss bursts, i.e., how losses are distributed in the flow. Several video sequences and loss patterns were employed to train the PSQA network. The obtained results manifested the impact of packet losses on the model's quality predictions. However, additional results showing the accuracy of the model in terms of correlation with subjective tests would be desirable. As stated by other authors, there are many other parameters which could cause the video quality degradation, so a comparison with other models that take into account other impairment factors would be necessary to confirm the good perspectives of this model.

In [154], the V-Factor model was presented. This methodology is based on the deep inspection of the bit-stream aiming to analyze if the measured impairments affect the video signal. The metrics employed by the V-Factor model are varied, namely, the content characteristics, compression and bandwidth constraints, or network-introduced impairments such as packet loss, delay, or jitter. The V-Factor quality prediction is obtained by the combination of the outputs of three main blocks: the bandwidth model, the complexity model, and the loss model. The former analyzes the bandwidth variation of the video stream. The complexity model makes an inspection of slices and macroblocks in order to analyze the variation of the quantizer and the motion vector. The latter block, together with the complexity model, evaluates the actual visibility of packet losses in the video signal. The final V-Factor value, which is the video MOS estimation, is computed by a codec-specific curve-fit employing the outputs of the three previous-stage blocks. Results are not conclusive because the work only showed how

the output MOS reacts to packet loss but no comparisons with other objective models or subjective tests are provided. Additionally, it would have been very interesting to know some deeper information about the algorithm insights, especially for the fitting-curve employed to gather the information extracted from the IP packet-headers and the video payload.

Joskowicz and Lopez presented a model [155] focused on the effects of frame-rate, encoding bit-rate, display size, and video content. The latter, which has an important weight in the user quality perception as discussed in previous sections, is derived from the average Sum of Average Differences (SAD) per pixel of the video-clip. SAD is an extended video-metric used for block comparison and for moving vectors calculations; therefore, the average SAD per pixel provides an overall estimation about the spatial-temporal activity of the entire video-clip, which in some way can be related to the content of the video. In order to obtain the quality estimation of the video under study, all the algorithm inputs are combined with nine fixed coefficients that are adjusted for each codec and additional video settings. As happened with packet-layer models, the need of employing empirical values seems to be the greatest drawback of this model, hindering its generalization.

In [156], a quality estimation methodology based on the model developed by Raake *et al.* [149] was presented. The novelty introduced by this proposal is that it includes a visibility framework, which integrates content-dependent features and employs a classification algorithm for the determination of the visibility of packet loss events. The video quality assessment model is similar to that employed by the T-V-Model, but the transmission-induced impairments factor incorporates the outcome of the visibility loss classifier used in a previous step. Results showed a great accuracy when detecting the visibility of packet losses. However, this model does not improve the correlation with subjective tests of that obtained by the T-V-Model, so the use of information extracted from the bit-stream is not exploited to reach better accuracy.

As stated in [158], the quality estimation models analyzed so far are quite simple approaches that can provide just some shallow quality predictions. The human evaluation process is much complex than these algorithms. Consequently, future developments of objective video quality measurement require algorithms with increasing complexity that need to be evaluated on large datasets in order to prove their performance and robustness in real world scenarios. Thus, **hybrid models** are the

closest approach to the human perception system, making use of information extracted from both, packet headers and the decoded video stream. Therefore, network-introduced impairments can be related with the video content, so a better understanding of the impact of the actual distortions on the perceived quality is attained. Because of the complexity inherent to these models, few works [159]–[163] addressing these methodologies can be found in the related bibliography in comparison with the approaches discussed above.

Yamagishi *et al.* presented a methodology accounting for the impairments introduced by the video compression process [159]. In this model, first a parameter extraction from the packets header and the video stream is done (e.g., bit-rate, packet loss, video content, etc.). Then, an estimation module for compression takes the bit-rate and the content-based information and outputs the video quality affected only by video-compression. Although a packet-loss estimation module was also described, authors stated that this was left for further study, and the distortion introduced by packet losses was not taken into account. Aiming to adjust each quality-estimation to the characteristics of the video content under evaluation, a series of subjective tests were conducted to extract some empirical parameters, employed in the estimation blocks. The accuracy of this model, in terms of Pearson Coefficient, RMSE, and Outliers Ratio, satisfies the minimum requirements established by the ITU-T; furthermore, an extension of this work could improve these results if the impact of other network-introduced impairments, as packet loss or delay, would be also considered.

In [160], a model based on an Adaptive Neural-Fuzzy Inference System (ANFIS) is described. Authors evaluated the impact of four parameters, namely, frame-rate, send bit-rate, packet error rate, and link bandwidth on the quality of three different-content video-scenes. Thus, three different Artificial Neural Networks (ANN) trained by the ANFIS were employed to assess the quality of the three video categories employed: slight, gentle, and rapid movement. Authors found that network level parameters like packet error rate or link bandwidth have a greater impact on video quality compared to application level parameters such as frame-rate and video send bit-rate. The main drawback found in this model is the need of subjectively pre-classify each video-sequence to one of the three categories described above in order to employ the appropriate ANN. Additional results comparing the correlation of the model with subjective test would be also recommendable in order to widen its applicability.

Sugimoto *et al.* presented a different approach in [161]. Their model accounts for the spatial and temporal image features extracted from the baseband signal and the quantizer-scale information obtained from the bit-stream. Each of these three features is obtained separately and then, they are integrated using the weighted Minkowski metric [164]. The weight coefficients were calculated empirically for a video test set. Results showed that the proposed method performs better than the PSNR metric, but just coding degradations are taken into account, ignoring important impairments introduced by the transport network. The use of additional coding schemes, bit-rates, etc. would also provide an idea of the model generalization capacity.

A hybrid model that deals with interactions between different-layer artifacts was presented in [162]. A first quality estimation provided by this methodology is composed by two metrics that estimate the strength of blockiness and blurriness artifacts. Then, a combination model is used to add the impact of the packet loss rate to the final quality estimation. Thus, the final hybrid quality score is obtained by the weighted combination of both quality estimator functions. The fitting factors employed were assigned to fixed values, and this is the main drawback of this algorithm, which works reasonably well under the controlled development conditions, but no other experiments were conducted in order to assess its accuracy with different video sequences affected by other impairments or artifacts.

Keimel *et al.* [163] presented a model based on a trilinear partial least square regression from a purely bit-stream feature-based metric combined with pixel-based features. Several video characteristics, e.g., slice type, Kbits per slice, or motion vector metrics, among others, were considered to obtain a first quality estimation based only on the bit-stream. In order to provide a pixel-based quality estimation, the pixel-level analysis block evaluates the following video characteristics: blockiness, blurriness, activity, predictability, motion continuity, edge continuity, and color continuity. Finally, both quality values are weightily combined to obtain the final output of the model, which is a quality prediction on a (0-1) scale. The main weakness of this model is the lack of flexibility since it was developed for just one particular codec and resolution. Additionally, the fixed fitting values may be inaccurate depending on the video content, given that some authors have demonstrated that different-content sequences cannot be analyzed in the same way because frame errors are more or less visible depending on the content characteristics [160].

2.4.2.2 Complete-user-experience Quality Assessment Models

In this section, the focus is on those models which include in their metrics additional parameters that allow a more complete user experience evaluation than a simple audiovisual quality assessment. As examples, although not focused just on IPTV, works in [19], [21], [165]–[167] provide a wide overview of the subjective-influencing factors that make the human quality-rate process a highly-difficult task to emulate. Authors of these works agree that there are complex interactions and dependencies among user, product, and context characteristics that are hardly reproducible by an objective quality assessment model. These works presented general taxonomies from different points of view and established an important baseline to build up quality estimation models from these perspectives.

Focused on IPTV services, Shin presented in [21] a general view of the great number of subjective parameters that impact on the user quality perception. In this work, a step forward was taken and numerical relationships between subjective parameters were provided. As shown in Fig. 2.13, the author considered that the perceived usefulness, quality of service, content quality, playfulness, and control impact on the user's attitude towards the IPTV service. Additionally, by considering the user's perception about the security and the cost of the service, Shin found a weighted relationship among all these service-features, which provides the user's intention to use IPTV.

Another different approach, in this case taking into account the user feedback is described in [168]. The proposed scheme, so-called Q-score, extracts a useful association between the inaccurate and indeterminately-delayed user's feedback and the different impairments introduced by the network. This action is performed through an off-line learning process by applying a statistical regression that provides the coefficients that quantitatively capture the relationship between the measured networks parameters and the customer's feedback. These coefficients are then employed in the on-line part of the model to accurately transform the measured network-objective metrics into a single quality index. Besides the real-time quality estimation, the Q-score was also proposed as a useful tool for other important tasks for the IPTV operators such as (i) the identification of significant Key Performance Indicators (KPI) affecting the QoE, (ii) the prediction of bad QoE, or (iii) the dimensioning of customer care workforce. The results revealed that the Q-score is able to predict 60% of service problems reported by customers with only 0.1% of false positive.

As discussed in previous sections, one of the most valued characteristic by the end user is the channel-zapping time. Different proposals focused on evaluating the relationship between channel-change time and the quality perceived by the user can be found in [169], [170]. Kooij *et al.* proposed a zapping time-to-MOS mapping in [169]. In this work, authors defined a minimum and a maximum zapping time to use a logarithmic interpolation between these extreme zapping times. The resultant function provides an accurately prediction (99% of correlation with subjective tests) of the MOS impacted by channel zapping time. In this work, a blank screen with the channel number was shown while channel zapping. On the other hand, the work in [170] evaluated channel zapping impact on the perceived quality when advertisements are displayed during zapping instead of the black screen usually employed. Godana *et al.* found that, for short zapping times (less than 0.65 s), customers prefer a blank screen than an advertisement. However, for medium and large zapping times the advertisements are preferred against the blank screen. Finally, analogous to [169], authors also proposed a model to estimate the MOS considering just the impact of the zapping time when advertisements are displayed during the channel-change time-interval. They obtained a very high correlation of 99% with subjective tests. Additionally to these works, Manzato *et al.* [171] and Ramos [172] presented brief reviews of different channel switching schemes aiming the reduction of zapping time for IPTV systems, considering also P2P architectures. As an example of these techniques, work in [173] proposed a dynamic channel-reordering according to the user preferences, including also a reordered delivery of time-shifted group-of-picture (GOP) sequences for the adjacent channels, in order to speed up switching latency. By using this technique, authors reduced channel change latency significantly, compared to traditional static reordering

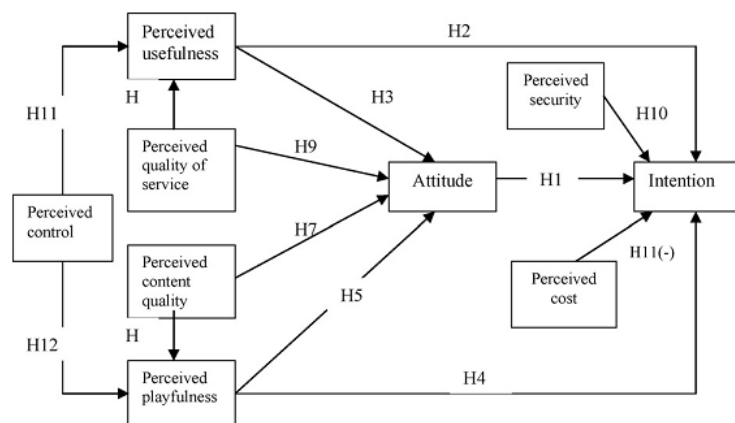


Figure 2.13. Multi-subjective-features-aware model. H items represent the different service's features perceived by users (extracted from [21]).

techniques. Another channel navigation scheme employing recent viewing data from all customers was presented in [174]; thus, this proposal identified hot and cold channels regarding the information extracted from the IPTV system, as well as the target viewer's own viewing history. In addition, a fast preview mode was provided for hot channels by prefetching them.

Another important factor that has a direct impact on customers' opinion about a service is their perception about the value of that service, i.e., if the service is worth its cost. Thus, Hayashi *et al.* [175] and Reichl *et al.* [176] developed different QoE assessment models regarding this factor. In the former [175], authors developed a methodology accounting for two main parameters: video quality (video coding rate) and pricing. By measuring the impact of each factor independently on the MOS, authors derived an expression to estimate the service value as a linear function of video quality and pricing on a psychological interval scale. Finally, they also obtained a MOS expression from the factor representing the value of the service. In turn, Reich *et al.* [176] developed a similar-concept model in which QoE is represented as a function of QoS and service cost. As the authors stated, the QoE depends on both the QoS offered by the provider and the customer expectations that are triggered by the corresponding tariff. Authors carried out an experiment where test users could turn the quality of a video sequence, with its corresponding price increase or decrease. Although this work did not provide any expression for calculating the service MOS from its quality and price (authors only gave a shallow description for this relationship), their results showed several customer-profiles with different behavior when tuning the quality, demonstrating the complexity of obtaining a general quality estimation model regarding the value of the service, as this is another factor that is particularly perceived by each individual user.

To sum up, in this section several approaches aiming to evaluate the quality of the IPTV service from a wider perspective has been reviewed. Due to the difficulties to simulate such internal human perception processes, there is still not a complete objective model gathering all the perceptual and psychological factors that a user experience when consuming multimedia contents. However, this should be the trend to follow in order to obtain accurate quality estimations that could be applied to improve the end user's satisfaction about future multimedia services.

2.5 Case Study: Evaluating the Quality of VoIP Services in a Real Environment

In this section, a real QoE subjective test is presented, aiming to serve as an exemplification of the different techniques discussed so far. This experiment, consisting in a comparative evaluation of two different coding schemes employed by two different VoIP applications, was carried out by means of both subjective and objective tests. The codecs were G.722 and Silk, and the VoIP applications Jitsi [177] and Skype [178]. This study was carried out in a controlled Ethernet scenario with 60 subjects. Each subject conducted a test in which two VoIP calls were established, one with Jitsi (codec G.722) and the other with Skype (codec Silk), resulting in a total of 120 VoIP calls for evaluation. The contribution is two-fold. First, at the end of each VoIP call, each subject answers a quality test to obtain her opinion about the service (subjective perspective - QoE). Second, all VoIP calls will be captured with a network analyzer tool at both communication endpoints, so that different network parameters can be monitored and analyzed (objective perspective - QoS). Thus, a comparison between the QoS and QoE results is carried out, obtaining a clear relationship between them.

2.5.1 Test-Bench and Components Features

The performance evaluation was carried out in the wired scenario shown in Fig. 2.14. The transmitter (TX) and the receiver (RX) were placed in two different locations separated by three hops (two intermediate nodes). These nodes had two different public IP addresses because of the need of Jitsi to access the Internet for its proper operation; whereas Skype allows the direct communication between TX and RX, Jitsi makes use of a re-transmitter located 18 hops away from the TX. In addition to the two mentioned VoIP applications, the packet-sniffer Wireshark [179] was also employed in order to capture the exchanged traffic between the two communication endpoints, allowing the analysis from a QoS perspective.

2.5.1.1 Locutions

As shown in Fig. 2.14, the tests consisted of a subject, located in the RX end-point of the communication, listening to a pre-recorded locution streamed from the TX side. Thus, the quality evaluation results were obtained in a MOS-LQ (Listening Quality) scale. The employed locutions were recorded in the Polimedia Studio at Universidad Polit cnica de Cartagena following the guidelines of the ITU-T Rec. P.830 [72]. The

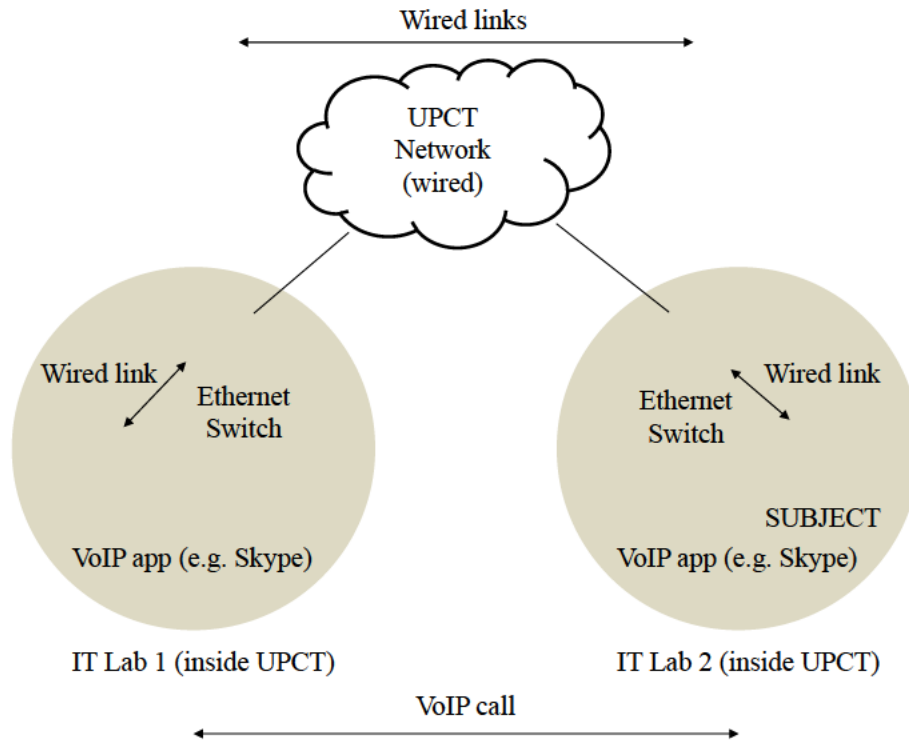


Figure 2.14. Employed test-bench (UPCT=Universidad Polit cnica de Cartagena).

recorded signals consisted on simple meaningful sentences, short (between 2 s and 3 s) and easy to understand. 4 different voices were selected (2 male and 2 female) and the sentences were pronounced smoothly and without emphasis in any part of them. The locutions did not present any diction deficiency or stuttering, and had a normalized volume.

To avoid any dependency between the obtained results and the individual characteristics of the voices, the masculine and feminine voices were evaluated separately. As aforementioned, two different male and two different female voices were used; thus, an equilibrated plan was obtained, reducing the risk of altered results because of the concrete features of a particular voice.

2.5.1.2 Transmitter and Receiver Equipment

The recordings were played in a laptop using Skype and Jitsi, which employ the Silk and G.722 codecs, respectively. In order to avoid environmental disturbances, the microphone on both sides was disabled, so that the only signal sent was the one recorded in the location. The main features of the TX computer are shown in Fig. 2.15 (a); it is noteworthy to mention that the sound card employed was a NVIDIA High Definition Audio. The RX system, where the test-subject was placed, was equipped with headphones with a frequency response of 18-20000 Hz, an efficiency of 115 dB

Información del sistema	Información del sistema
Fecha y hora actuales: martes, 02 de agosto de 2011, 19:27:10	Fecha y hora actuales: miércoles, 03 de agosto de 2011, 0:57:18
Nombre del equipo: SALVA-PC	Nombre del equipo: ANTONIO-PC
Sistema operativo: Windows 7 Home Premium 64 bits (6.1, compilación 7601)	Sistema operativo: Windows 7 Home Basic 64 bits (6.1, compilación 7600)
Idioma: español (configuración regional: español)	Idioma: español (configuración regional: español)
Fabricante del sistema: Hewlett-Packard	Fabricante del sistema: Hewlett-Packard
Modelo del sistema: HP Pavilion dv4 Notebook PC	Modelo del sistema: HP G62 Notebook PC
BIOS: Default System BIOS	BIOS: Default System BIOS
Procesador: Intel(R) Core(TM)2 Duo CPU P8400 @ 2.26GHz (2 CPUs), ~2.3GHz	Procesador: Intel(R) Core(TM) i3 CPU M 330 @ 2.13GHz (4 CPUs), ~2.1GHz
Memoria: 4096MB RAM	Memoria: 4096MB RAM
Archivo de paginación: 1979MB usados, 6205MB disponibles	Archivo de paginación: 1947MB usados, 5966MB disponibles
Versión de DirectX: DirectX 11	Versión de DirectX: DirectX 11

(a) Transmitter PC characteristics

(b) Receiver PC characteristics

Figure 2.15. Equipment characteristics.

SPL/V, a maximum input power of 200 mW, an impedance of 32 Ohms, and a wire of 2.5 m (99.99% OFC, unilateral). Fig. 2.15 (b) presents the most relevant features of the computer employed by the RX; in this case, the soundcard used was a Realtek High Definition Audio.

2.5.1.3 Participants

The subjects selected to carry out the quality evaluation were randomly elected from a population used to employ the telephone service. However, they accomplish the following conditions:

- i) Participants have not participated in previous telephonic circuits quality tests or similar.
- ii) Participants have not collaborated in previous subjective test, of any kind, during the last 6 months, or listening quality tests, during the last year.
- iii) Participants have not previously heard the employed locutions.

Further description about the participant election process can be found in [2].

2.5.1.4 Experimental Methodology

During the design stage of the experience, it was also taken into account the playing order of the selected recordings aiming to avoid any effect caused by this fact. The 60 participants were divided into groups of 20 subjects, which listened to the locutions in the same order; thus, the first group firstly listened to the 2 male locutions and then to the 2 female ones; the second group of participants firstly listened to the 2 female recordings and then to the 2 male locutions and, finally, the last group listened to the recordings in an alternate order. The combined length of the four recordings was 5 min 17 s.

For a given group of subjects, the test duration should be short enough to avoid the onset of fatigue. In this test, each participant listened to her assigned group of locutions twice: firstly to evaluate the Silk codec employed by Skype and, then, to score the G.722 codec used by Jitsi. Therefore, each session had a maximum length of 15 min.

2.5.2 QoE Data Collection

2.5.2.1 Questionnaire

As described in the ITU-T Rec. P.800 [2], it was necessary to develop a questionnaire that subjects answered separately for each codec/application under study. By means of the provided answers, it was extracted the subjective quality perception that the participants had about each evaluated coding scheme (and thus, about the VoIP application). The key points considered in the elaborated questionnaire were:

- The voice quality appreciation in the Absolute Category Rating (ACR) scale (1 – 5).
- Issues to hear through the network link.
- Effort made to understand the locutions meaning.
- Best clarity of the recordings according to its nature: male or female.
- The subjects' Internet and VoIP frequency of use.

2.5.2.2 Calculations

In the following, the most relevant mathematical apparatus employed to analyze the collected subjective data are presented [180].

The skewness is a measure of the asymmetry of a probability distribution about its mean. If the skewness is equal to 0, the distribution is symmetric. Negative skew indicates that the tail on the left side of the probability density function is longer or fatter than the right side, and vice versa. Skewness is calculated as formula (2.10) indicates,

$$\alpha_3 = \frac{\sum (x_i - \bar{x})^3}{(N-1) S^3} \quad (2.10)$$

where x_i is the value of each data point, \bar{x} is the mean, N is the number of data points, and S is the standard deviation of the sample.

The kurtosis is the measure of the peak of a distribution, and indicates how high the distribution is around the mean. Positive values of the kurtosis indicate that the distribution has a more acute peak around the mean; in turn, a negative kurtosis means a wider peak of the distribution. It is calculated as shown in expression (2.11),

$$\alpha_4 = \frac{\sum (x_i - \bar{x})^4}{(N-1) S^4} \quad (2.11)$$

where x_i is the value of each data point, \bar{x} is the mean, N is the number of data points, and S is the standard deviation of the sample.

The ANOVA (Analysis of Variance) is a statistical technique for determining the degree of difference or similarity between two or more groups of data under different treatments. It is based on the comparison of the average value of a common metric variable calculated for the different groups determined by an independent and nominal variable, so-called factor. The null hypothesis is considering that all groups have the same average, i.e., there are no statistical significant differences among group averages. Rejecting the null hypothesis implies that different treatments result in altered effects.

2.5.3 QoS Data Collection

The collection of network-metrics was carried out by using the packet sniffer and traffic analyzer tool Wireshark. By means of this application, all traffic exchanged between TX and RX was captured for its further analysis. As each participant had to listen the locutions twice (once per VoIP application), and the capture was done in both extremes of the communication, 4 files were collected for each subject, generating a total amount of 240 captured files.

Regarding the QoS analysis, it has been done taken into account the following metrics:

- Total employed bandwidth
- Packets size
- Jitter
- Packet Loss Ratio (PLR)
- Bandwidth consumed by other non-VoIP applications

All these metrics, except the latter, were measured by filtering the capture files, discarding all the packets unrelated to the VoIP conversation. On the other hand, by means of the opposite filter, it was analyzed the traffic generated by other non-VoIP applications or services running in the employed computers.

In order to ease the filtering task, different *awk* [181] scripts were created. Thus, the capture files generated by Wireshark (.pcap) were converted to plain-text files (.txt) to obtain the maximum compatibility with *awk*. Concretely, 3 different scripts were developed: one for the captures in the TX side, another one to analyze the captures obtained in the RX side, and, finally, a third script was employed to evaluate the non-VoIP traffic.

2.5.4 QoE Results

Fig. 2.16 and Fig. 2.17 show the ratio of participants per gender and the Internet and VoIP applications frequency of use, respectively. From the total of subjects, 76.67% were male and 23.33% were female. In a scale from 1 to 5, the participants determined their self-use frequency of 2.60 for VoIP applications and 4.31 for the Internet. Clearly, the Internet use was high and widespread among the participants, but they were not used to VoIP applications. This benefits the subjective study since the participants did not present prior experiences and, consequently, expectations about the applications quality.

Regarding quality perception, the percentage of participants that felt issues when listening to the locutions for both VoIP applications is shown in Fig. 2.18. In both cases, the percentage of subjects that did not suffer any difficulty to hear the recordings was greater than that representing the subjects with difficulties. Nevertheless, this difference was more noticeable in the case of Jitsi (78% vs. 22%) than in the case of Skype (68% vs. 32%).

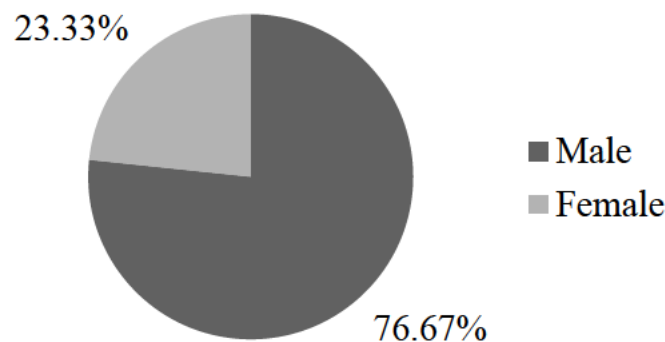


Figure 2.16. Participants' gender.

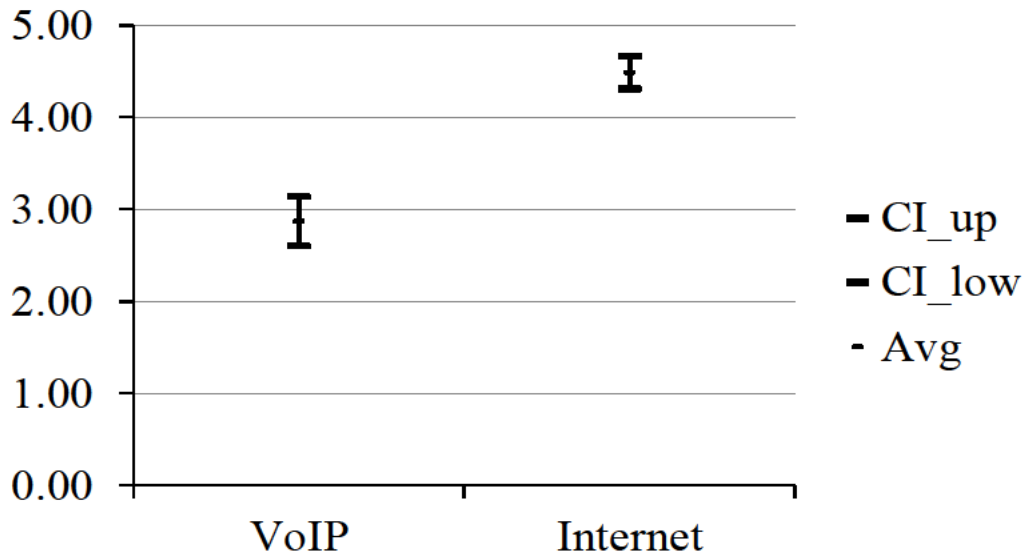


Figure 2.17. Participants' Internet and VoIP applications frequency of use in a scale 1-5 (CI=Confidence Interval).

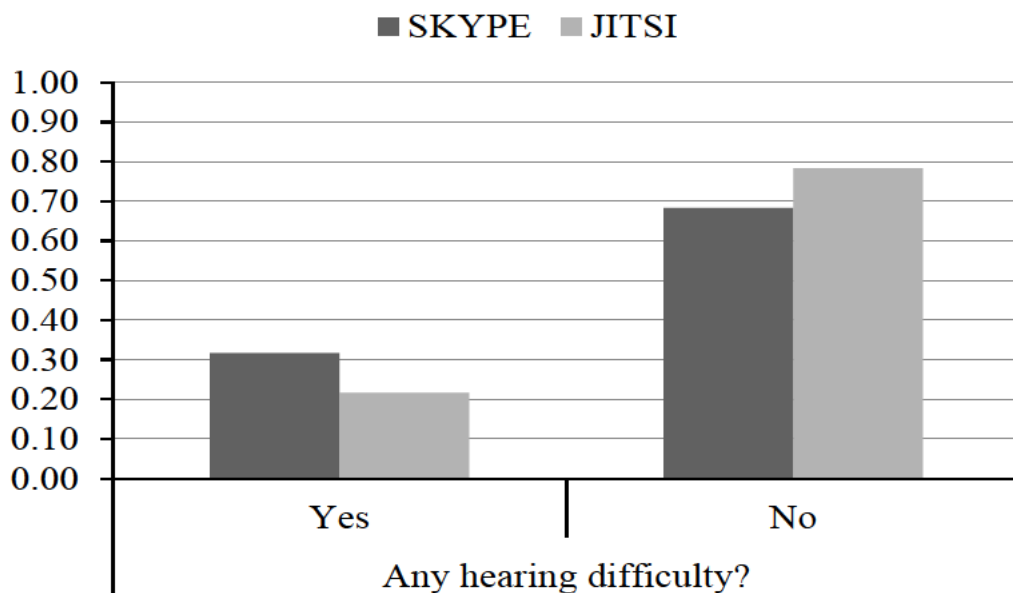


Figure 2.18. Ratio of participants that had hearing difficulties (self-evaluation).

Fig. 2.19 and Fig. 2.20 depict a comparison among the different impairments perceived by the subjects when employing both applications. Using Skype (Fig. 2.19), an annoying noise or buzz was the most recurrent impairment, noticed by the 52.63% of the participants, followed by the voice distortion and the complete disconnection of the communication with similar values. Finally, the least perceived errors were low volume, short disruptions, or volume changes with 5.26%. In the case of Skype, neither echo nor crosstalk was noticed. Regarding Jitsi (Fig. 2.20), there were two relevant impairments, annoying noise and buzz, and volume variations and short disruptions, in both cases

with approximately 38,46% of the participants detecting them. No cases of crosstalk, low volume, or echo were perceived. The only impairment with similar levels of perception in both applications was distortion, with 21.05% of subjects detecting it in Skype, and 15.38% in Jitsi.

Fig. 2.21 shows a comparison between the voice clarity perceived by the subjects depending on the speaker's gender. In both applications, there was not a clear distinction between the clarity provided by both genders, revealed by the similar percentages obtained for male and female voices quality. In Skype, 10% of the

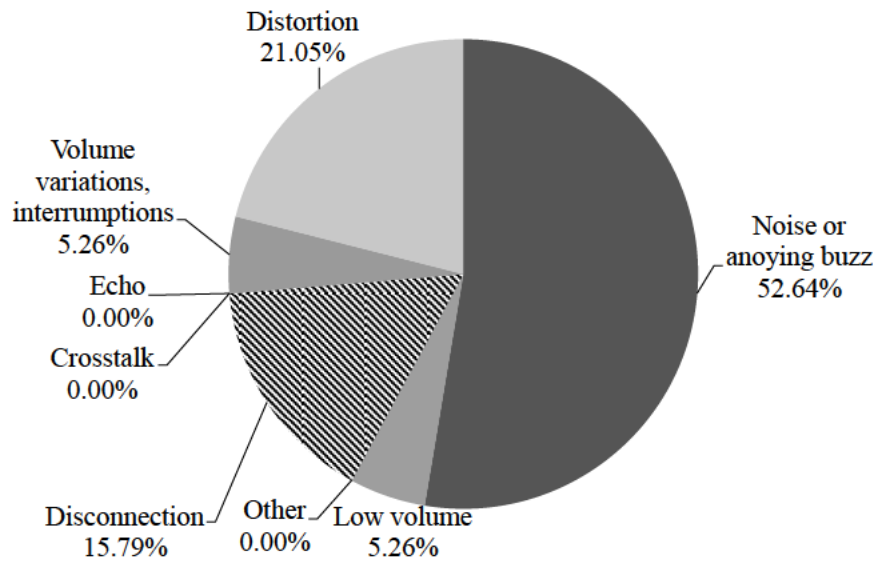


Figure 2.19. Types of hearing impairments perceived using Skype.

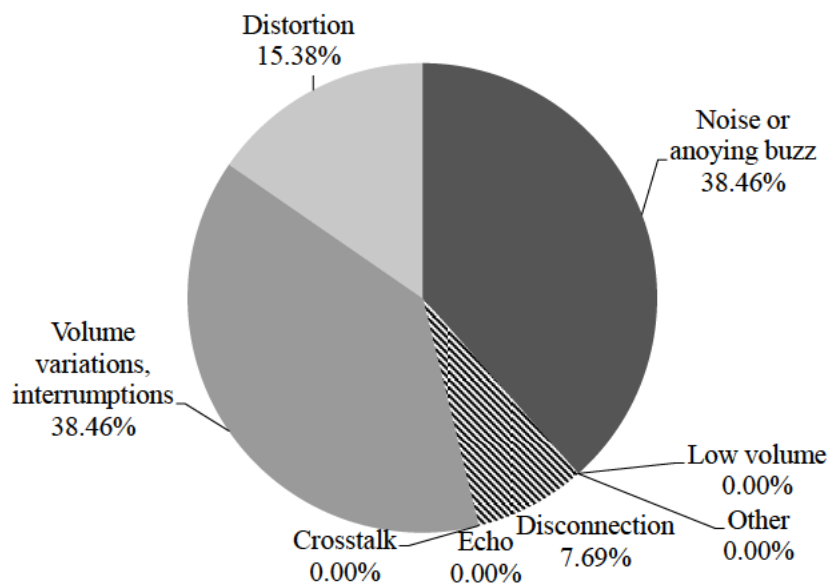


Figure 2.20. Types of hearing impairments perceived using Jitsi.

participants perceived the male voices with more clarity, against 20% of subjects, who rated the female voices with greater scores. The majority of participants (70%) did not find any quality difference between masculine and feminine voices. In Jitsi, 7% of the subjects found clearer the male voices, against 18%, who rated the female voices with greater levels of quality. The remaining 75% made no distinction between both gender voices regarding their clarity.

Fig. 2.22 and Fig. 2.23 depict the obtained listening quality (MOS) and the necessary understanding effort, respectively. It is clear that the understanding effort was similar no matter the employed application (see Fig. 2.23). In turn, the MOS-LQ obtained for Jitsi was greater than that attained for Skype, with average values of 4.48 and 3.92, respectively, as presented in Fig. 2.22.

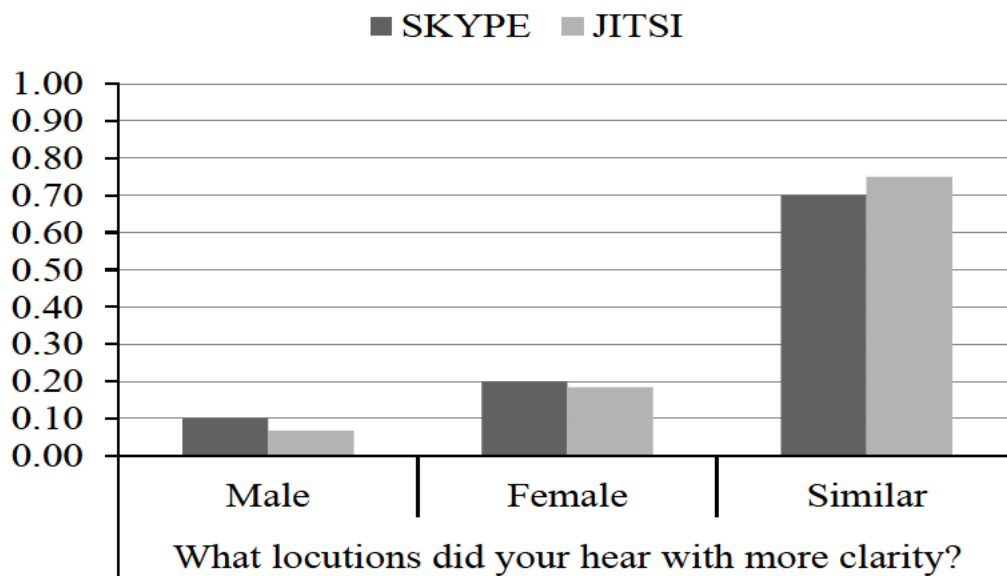


Figure 2.21. Locutions with more clarity by voice gender.

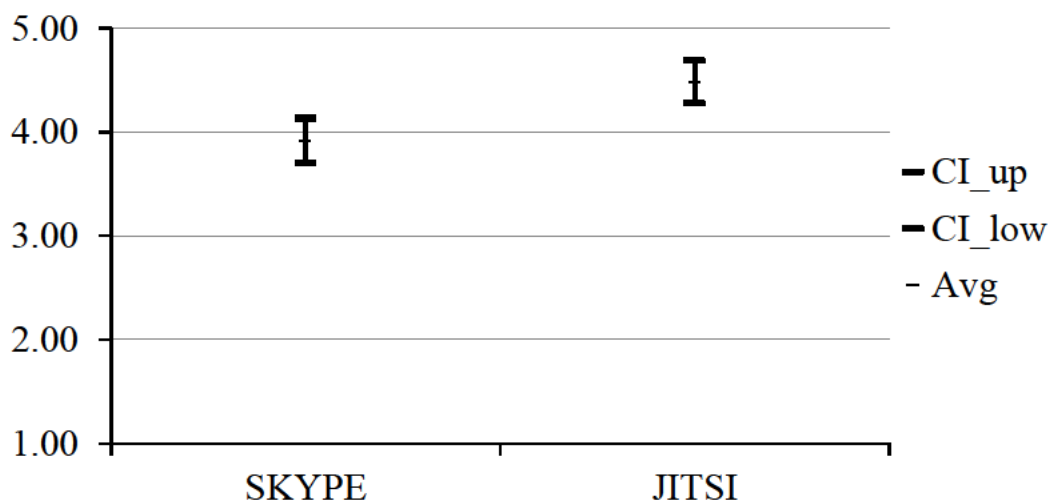


Figure 2.22. Voice quality in MOS scale (CI=Confidence Interval).

Fig. 2.24 shows the trend over the average that followed the voice quality in MOS scale, the frequency of Internet use, and the frequency of VoIP applications use, by evaluating the skewness of their corresponding distributions. The clearest difference between both applications was observed in the distribution of the voice quality in MOS scale. Although in both cases the skewness of the voice quality (MOS) was negative, which entails a trend toward higher values than the corresponding average, the absolute value in the case of Jitsi was noticeably greater than in the case of Skype. This means that the participants' perception towards Jitsi was even more positive than the average results showed (see Fig. 2.22). For the rest of parameters, the same skewness was

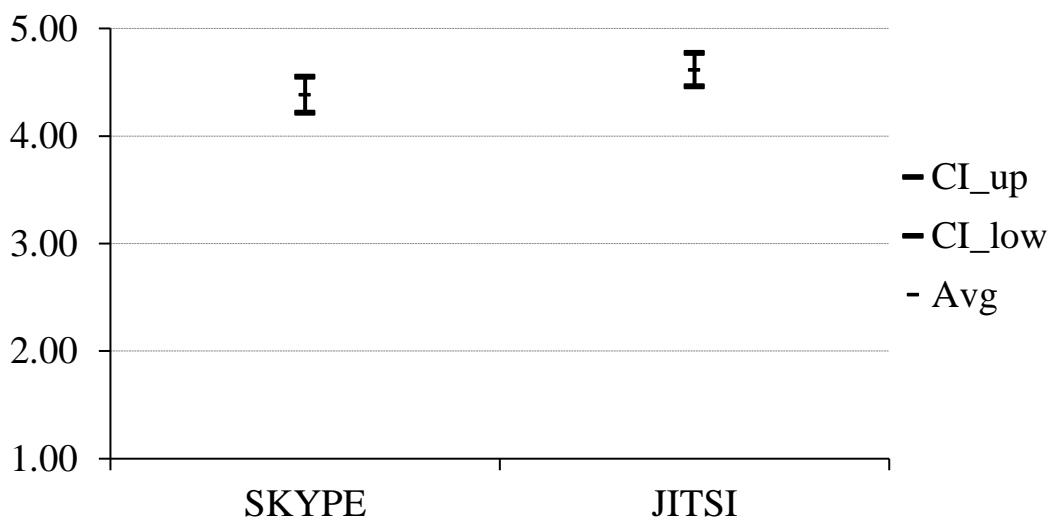


Figure 2.23. Understanding effort (CI=Confidence Interval).

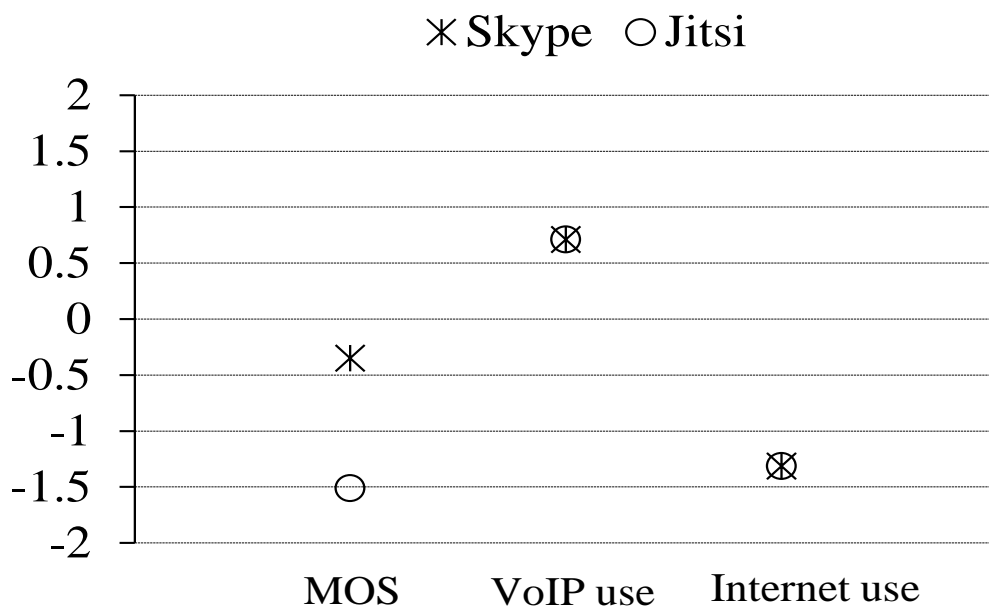


Figure 2.24. Skewness (observe that Skype and Jitsi get the same skewness for frequency of use of VoIP and the Internet).

TABLE 2.8. One-way ANOVA results

	Skype		Jitsi	
	F	p	F	p
Voice Quality in MOS Scale vs. Gender	1.93	0.170	0.21	0.647
Voice Quality in MOS Scale vs. Frequency of Internet Use	0.75	0.528	2.34	0.083
Voice Quality in MOS Scale vs. Frequency of VoIP Use	1.04	0.396	0.41	0.799
Frequency of VoIP Use vs. Gender	1.41	0.239	1.41	0.239
Frequency of VoIP Use vs. Frequency of Internet Use	2.06	0.116	2.06	0.116
Frequency of Internet Use vs. Gender	2.02	0.161	2.02	0.160
Frequency of Internet Use vs. Frequency of VoIP Use	3.23	0.019	3.23	0.019

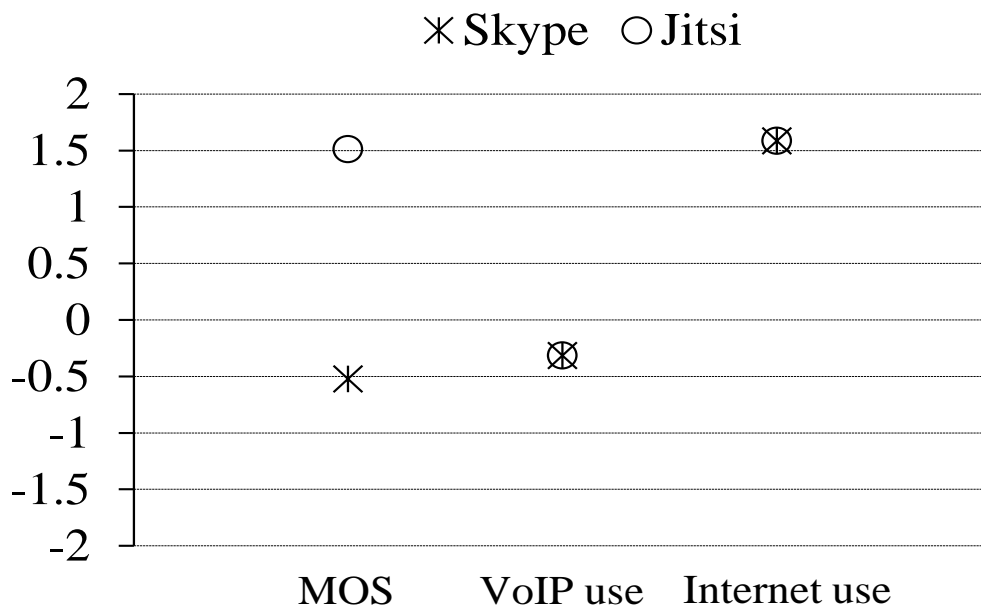


Figure 2.25. Kurtosis (observe that Skype and Jitsi get the same kurtosis for frequency of use of VoIP and Internet).

obtained for both VoIP applications. In the case of the VoIP applications usage, it was attained a positive skewness around 0.6. Thus, the distribution trended to lower values than the average, corroborating that the participants were even less used to VoIP applications than the average value showed in Fig. 2.17. The opposite behavior was observed for the obtained skewness related to the frequency of Internet use.

Fig. 2.25 shows the obtained kurtosis for the distributions of the same metrics previously discussed, namely, the voice quality in MOS scale and the frequency of use

of the Internet and VoIP applications. As in the skewness analysis, the most relevant difference between VoIP applications, regarding the kurtosis analysis, was attained for the voice quality in MOS scale. Kurtosis measures how outlier-prone a distribution is. The kurtosis of the voice quality (MOS) had a negative value when using Skype, which entails a flat distribution for this group of data. However, the kurtosis of the voice quality (MOS) was positive for Jitsi, so the distribution shape trends to present a peak compared with the normal distribution. This behavior can be explained in two different ways. It could be a sign of outliers, meaning that a few of the respondents gave low (high) ratings whereas the majority of the respondents agreed on a higher (or lower) quality. Second, it could be the result of a different situation experienced by the user because of a particular application misbehavior. Therefore, either some participants suffered volume variations and short disruptions (one of the most common impairment observed in Jitsi as shown in Fig. 2.20), or there were some outliers. For the rest of parameters they showed the same levels of kurtosis, being negative (near to 0) in the case of the frequency of use of VoIP applications and positive for the frequency of Internet use.

The results for one-factor ANOVA analysis (ANOVA One-Way) conducted on each factor under study are presented in Table 2.8. This tool provides insight into whether a factor impacts on the average response of another one. In this study, a confidence interval equal to 95% ($\alpha = 0.05$) has been used. If the value of p is lower than α , the null hypothesis is rejected, and the alternative is accepted, and vice-versa. The null hypothesis is that there are no significant statistical differences between population averages. Taking all these considerations into account, in almost all of the results showed in Table 2.8 the value of p is greater than α ; for that reason the following conclusions can be extracted. The participants' gender and frequency of Internet and VoIP use did not impact on the overall appreciation (voice quality in MOS scale) that they had about each VoIP application. The participants' frequency of VoIP use neither affected the gender nor the participants' frequency of Internet use. Regarding the frequency of Internet use, it did not impact on the participants' gender, but it did on the frequency of VoIP use. This is represented by a p value of 0.019 for both applications, lower than α and, therefore, making the null hypothesis to be rejected. Therefore, it is concluded that the participant's quality perception was not affected by the other factors considered, such as previous experiences, which might lead to altered results.

2.5.5 QoS Results

2.5.5.1 Bandwidth

There is a straight relationship between the bandwidth consumed by a multimedia service and the provided quality. However, in a shared medium it is desirable to employ as less bandwidth as possible so that other services also run properly. Fig. 2.26 shows the bandwidth consumed by both applications under study. It is clear that Jitsi needs greater bandwidth than Skype. Aiming at providing deeper analysis of this result, Fig. 2.27 and Fig. 2.28 depict the evolution of the bandwidth consumption during 5 different VoIP transmissions for both applications. These calls have been selected randomly from

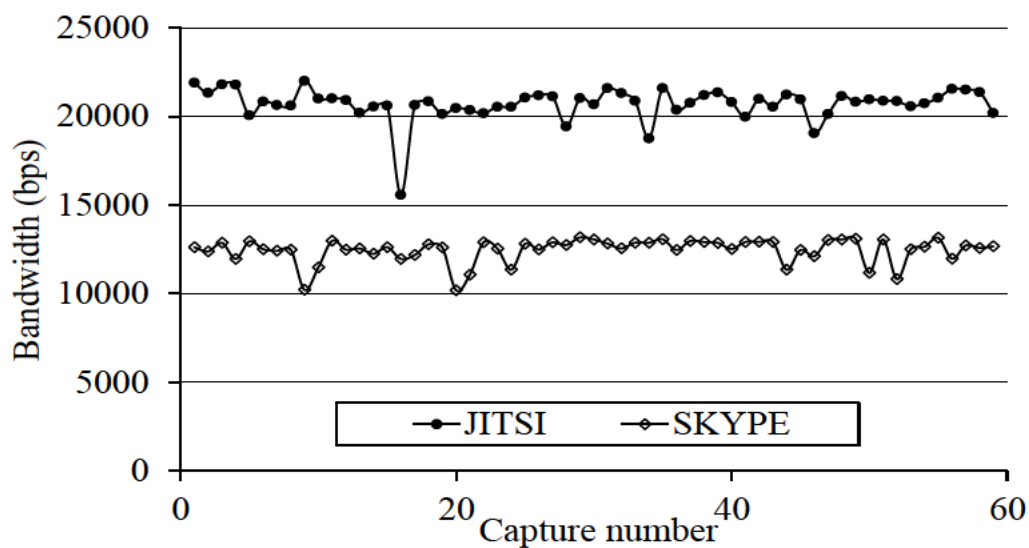


Figure 2.26. Overall consumed bandwidth in the RX side. Results for Skype and Jitsi.

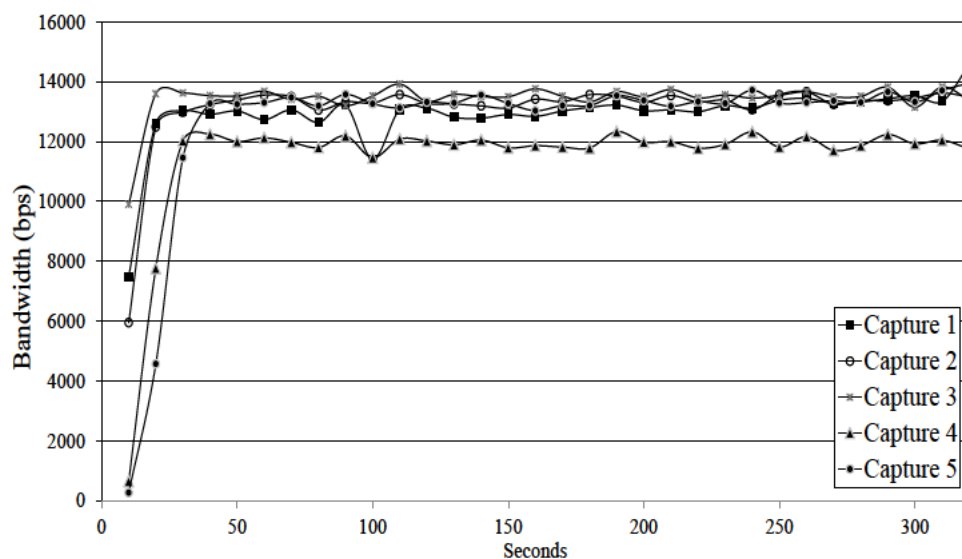


Figure 2.27. Detailed bandwidth consumption in the RX side. Results for Skype.

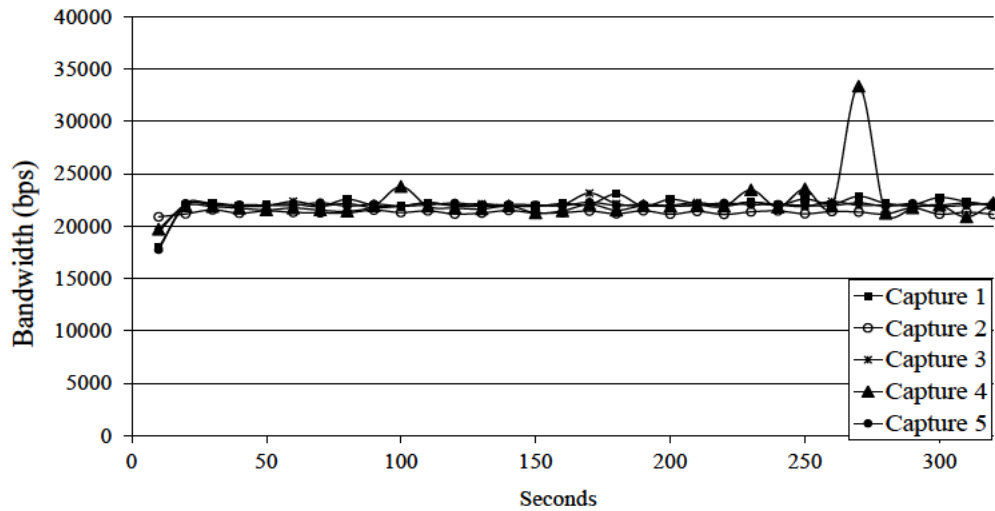


Figure 2.28. Detailed bandwidth consumption in the RX side. Results for Jitsi.

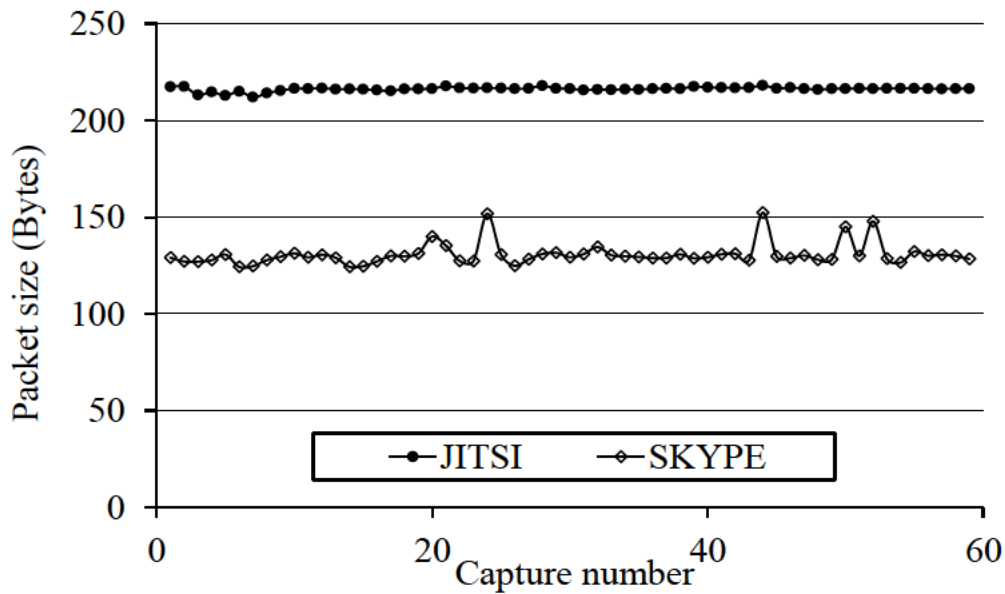


Figure 2.29. Average packet size.

the 60 monitored communications. The showed bandwidth values have been calculated in time-intervals of 10 s, from the beginning until the end of the transmission. Observe how Skype employs less bandwidth at the beginning of the call, and it is growing during the first 40 s until its stabilization at 13000 bps. On the other hand, the bandwidth employed by Jitsi is more stable, and its stabilization is attained after the first 25 s at 21800 bps.

2.5.5.2 Packet Size

A comparison between the sizes of the packets employed by both applications is shown in Fig. 2.29. It is clear that Jitsi makes use of larger packets than Skype. Jitsi uses

packets varying between 213 B and 217 B, and Skype generates packets between 124 B and 152 B in length. These intervals manifest that the packets size is more stable for the case of Jitsi.

2.5.5.3 Jitter

In order to provide good levels of quality in VoIP services, the delay between packet arrivals should be as low and non-variable as possible. Otherwise, users would perceive a lack of fluency or even voice-cuts.

The employed packetization interval was 20 ms of voice per packet, so that, the delay between arrivals must be lower than that value to ensure high values of QoE, avoiding noticeable impairments in the received signal. Moreover, the variance among different-packet arrival delays (jitter) should be as small as possible to avoid the perception of the voice packetization process.

Fig. 2.30 presents the average jitter obtained for each VoIP transmission employing the two applications under study. Observe that Skype presents more jitter variation than Jitsi, but most values are quite similar, around 10 – 11 ms. Furthermore, the greatest value for jitter is 13 ms, thus fulfilling the aforementioned delay and jitter requirements. Finally, Fig. 2.31 shows the jitter standard deviation for both, Skype and Jitsi. In average, Jitsi presents more variation in the jitter than Skype.

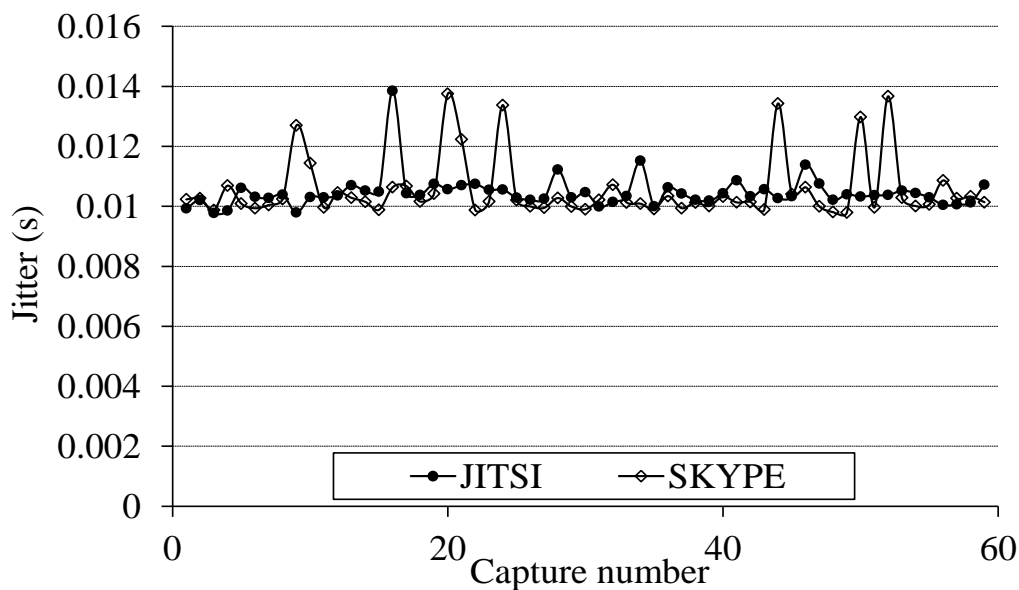


Figure 2.30. Jitter average.

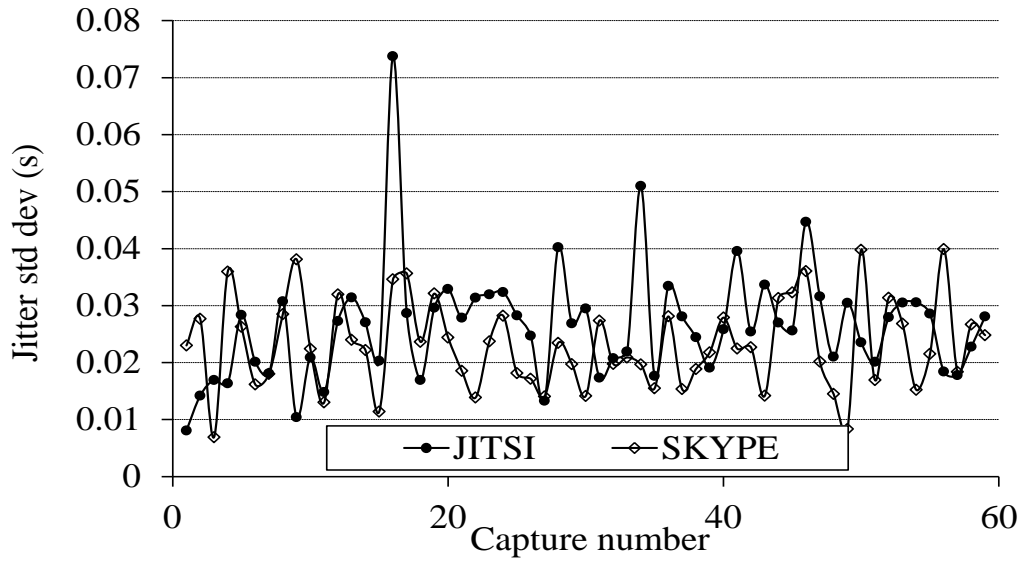


Figure 2.31. Jitter standard deviation.

2.5.5.4 Packet Loss

The main challenge that VoIP services have to face is their tough temporal requirements and the use of UDP as transport protocol. This protocol is low tolerant to packet loss due to its incapability to retransmit packets. Therefore, the loss of irrecoverable packets causes a deep decrease in the conversation quality. The ITU-T established the maximum PLR threshold to provide VoIP services with acceptable levels of quality in the Rec. G.114. This threshold was set to a maximum PLR of 5%; over that value, the VoIP QoE suffers a dramatic drop.

As shown in Fig. 2.32, the attained PLR during the study do not overstep the mentioned threshold, obtaining maximum PLR values of 3.01% for Skype and 1.97% for Jitsi.

2.5.5.5 Non-VoIP Bandwidth

In a regular network, in which several nodes share the access medium, it is usual that other applications generate non-VoIP packets. This traffic is so-called background traffic and could have a severe impact on the overall performance of the voice service. As shown in Fig. 2.33, the bandwidth occupied by other applications different from those under study was quite low, with the exception of some traffic peaks detected during Skype testing.

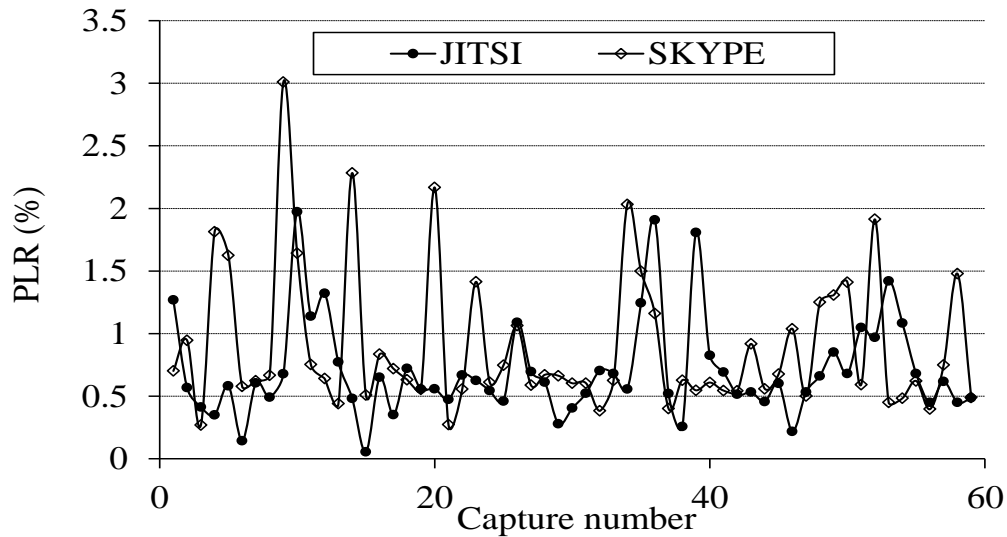


Figure 2.32. Packet loss rate.

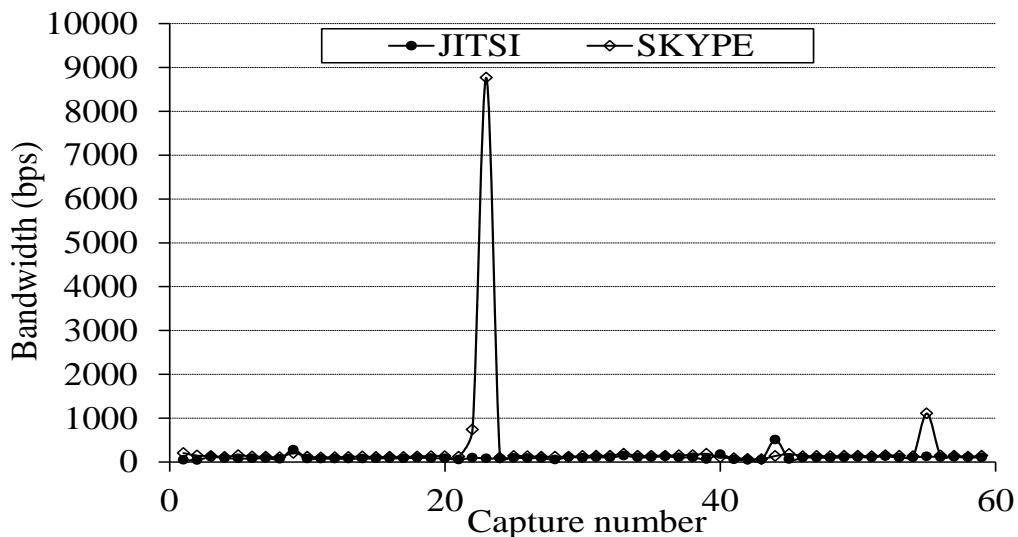


Figure 2.33. Bandwidth consumed by non-VoIP traffic.

2.5.5.6 Delay

As aforementioned, the delay is another key factor to keep under control in order to provide the desired level of quality. A high delay could cause echo or voices superposition, which represents annoying disturbances to the talkers. The ITU-T has established a maximum Round Trip Time (RTT) of 300 ms in order to avoid the impairments caused by the delay introduced by the network (ITU-T Rec. G.114 [182]). Before the study, a RTT analysis was done by using the ping utility from both extremes of the communication. The obtained results are depicted in Fig. 2.34. The delay obtained is between 15 ms and 17 ms, values noticeably lower than the established threshold of 300 ms, so the quality of the service is not affected by the impact of the delay introduced by the network.

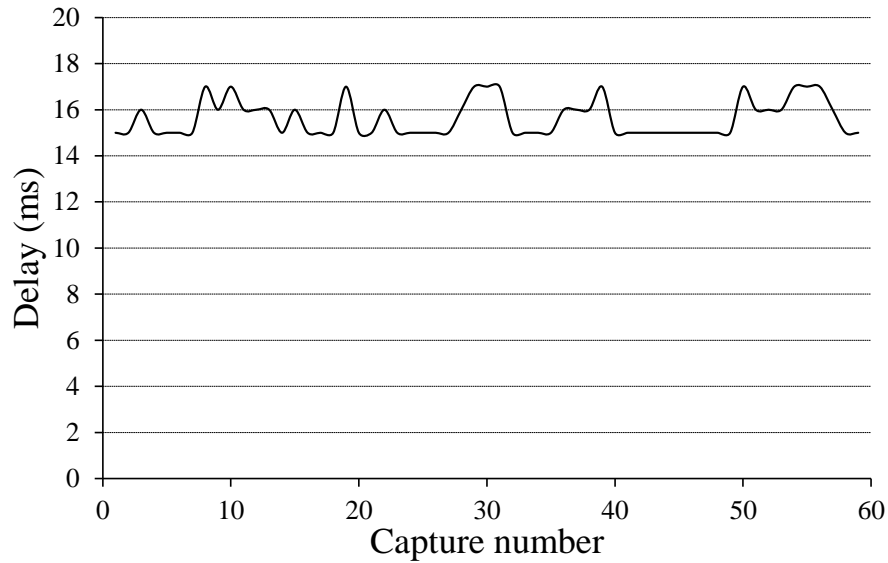


Figure 2.34. Network delay.

To sum up, this section described the process to evaluate the quality provided by multimedia applications from both objective and subjective perspectives. The objective results showed similar performance for Skype and Jitsi in terms of packet loss and jitter, which a priori are the most influent parameters on the quality. However, from the subjective analysis, it has been extracted that the best quality perceived by the end-users is achieved with Jitsi. As it is confirmed in this chapter, the only relevant objective difference between Jitsi (codec G.722) and Skype (codec SILK) is that the former makes use of bigger packets than Skype, consequently consuming more bandwidth. Thus, an important conclusion is that the more the bandwidth used by the voice codec, the better results in terms of MOS. For that reason, there exists a clear trade-off between bandwidth consumption and quality perceived by the end-user.

As general conclusion, it has been demonstrated the subjacent relationship between QoS and QoE. In scenarios like the one analyzed in this chapter, in which the impact of the network-introduced impairments is negligible, other parameters gain notoriety. In this case, system configuration parameters, such as packet size and bandwidth consumed by each codec (application), have made Jitsi to receive the best rating by the users. These factors are not usually studied as other more popular parameters such as PLR and delay. Consequently, it is important to remark that every objective parameter in the network should be considered when evaluating the quality of multimedia services by using objective metrics.

2.6 Conclusion

This chapter has provided the necessary background, regarding the QoE concept, for the further understanding of the following chapters. First, a series of QoE definitions were provided, discussing about the importance of this quality estimation paradigm in current multimedia services. It was also reviewed the standardization-bodies efforts in order to clarify and homogenize what should be understood when we refer to the term QoE. Then, an extensive survey of current methodologies employed to estimate QoE in VoIP services was presented. It was shown a detailed review of the current standards and non-standard proposals related with objective non-intrusive VoIP QoE estimation. The focus was on these methodologies, instead of the reference-based models, because obtaining QoE estimations at real-time assists network operators and service providers in performing efficient QoE management. Detecting and correcting quality drops at real time is one of the keys for the VoIP user's satisfaction, which does not have to wait until the end of the call to perceive the increase on the service quality.

Thereafter, a comprehensive survey on QoE assessment methods for Future TV services was presented. Similar to the case of VoIP QoE estimation models, this study reviewed separately the standard methodologies and the non-standardized proposals. Despite the great amount of models aiming at estimating the QoE in terms of MOS of diverse video services, there is still not a general model applicable for the great variety of complex conditions regarding video transmission, namely, coding scheme, bit-rate, resolution, service's architecture (live, on-demand...), etc.

Finally, a real experiment regarding the QoE evaluation of two real VoIP services was presented. The process to evaluate the quality provided by multimedia applications from both objective and subjective perspectives was described. From the obtained results, it can be extracted the existence of relationship between QoS and QoE. In scenarios with low level of network-introduced impairments such as the one analyzed in this chapter, other parameters, different from the typically analyzed in quality evaluations, gain notoriety. For that reason, it is important to remark that every objective parameter in the network should be taken into account when evaluating the QoE of multimedia systems from network metrics.

2.7 Contributions Related to this Chapter

R. Sanchez-Iborra, M.-D. Cano, S. Moreno-Urrea, and J. J. P. C. Rodrigues, “QoE measurements and analysis for VoIP services,” in *Emerging Research on Networked Multimedia Communication Systems*, D. Kanellopoulos, Ed. IGI Global, 2015, pp. 285–308.

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “Revisiting VoIP QoE assessment methods: are they suitable for VoLTE?,” *Computers Standards & Interfaces*. Under review. (I. F. 0.879).

R. Sanchez-Iborra and M.-D. Cano, “A comprehensive survey on Quality of user Experience assessment methods for future TV services,” *ACM Computing Surveys*. Under review. (I. F.: 3.373).

Chapter 3

Performance Evaluation of 802.11g Infrastructure Networks Supporting VoIP Traffic: Impact of the PHY Layer and Voice Codec on QoE

3.1 Introduction

The use of multimedia services such as VoIP, video streaming, IPTV, etc., over wireless networks is more and more usual, as discussed in the previous chapter. On the one hand, users expect the same quality levels when using these multimedia applications disregarding the technology (wired or wireless) in use. On the other hand, multimedia traffic over wireless networks suffers particular performance degradation due to the inherent characteristics of the transmission media that is not present in wired communications. Despite the considerable work done in the related literature about VoIP performance in wireless networks, even considering QoS and/or QoE, to the best of author's knowledge there are few published works addressing the detailed effect of the physical layer on QoE in VoIP traffic. For instance, the physical layer is influenced by path loss, shadowing, and fading, hence affecting the QoE as well.

Thus, in this chapter the effect of fading channels on VoIP transmissions over wireless networks is studied. Specifically, the focus is on wireless local area networks (IEEE 802.11g) in infrastructure mode. As great part of networking research experiments are usually performed in simulation platforms, it is important modeling the transmission channel as realistic as possible to obtain accurate results. For that reason, a series of analytical algorithms has been developed aiming to predict the power signal level at any point in the space, by considering the impact of the wireless transmission medium on the transmitted signal. Thus, in this chapter it is evaluated the influence of different propagation models (Rayleigh [183], Rice [184], Nakagami- m [185], and Free Space [186]) on the quality of VoIP calls in terms of: (i) maximum system capacity in terms of valid simultaneous VoIP calls in the network, (ii) evolution of the system capacity and the QoE (in terms of MOS) of the calls as a function of the distance from

the wireless terminal to the access point, and (iii) average delay and packet losses in the VoIP calls. In addition, the performance of two different codecs (G.711 A-law at 64 Kbps and G.726 at 24, 32, or 40 Kbps), with two packetization intervals (10 ms and 20 ms), and two different available wireless bandwidths (11 Mbps and 54 Mbps) is compared.

Depending on the environment under study (outdoor, indoor, industrial, etc.) different propagation models should be used to accurately characterize the transmission medium. This is because each propagation model presents its own characteristics and particularities. For that reason, in the following, a brief review about the propagation models employed in this thesis to characterize the wireless physical layer is presented. Then, the performance of IEEE 802.11g networks configured in infrastructure mode, supporting highly demanding traffic such as VoIP is explored. To this end, the influence on the quality of the voice-calls of the wireless physical layer and the voice coding scheme employed is evaluated.

3.2 Propagation Models

As mentioned above, it is necessary a formal characterization of the wireless physical layer in order to obtain realistic results when studying wireless networks from both analytical or simulation perspectives. This is achieved by means of mathematical propagation models. Given the high complexity of characterizing the wireless transmission medium due to its random nature, it has been developed several probabilistic propagation models that simplify this task [183]–[185]. However, there are other approaches such as deterministic or geometric models, e.g., ray-tracing models [187], or other based in empirical measurements, e.g., Okumura-Hata model [188].

Thus, selecting the more proper propagation model, depending on the evaluated environment and the precision requirements is crucial. In this thesis, different probabilistic models have been chosen due to its low computational requirements, compared to those needed by deterministic models. Moreover, the probabilistic models under consideration allow simulating different scenarios, taking into account the effect of the fading channels on the transmitted signals. Specifically, the propagation models used during the experimental-work conducted in simulation platforms were: the Rayleigh model [183], the Rice model [184], and the Nakagami- m model [185]. The former permits to simulate an indoor situation when there is no line-of-sight between

transmitter and receiver. The Rice model simulates an indoor simulation when there is line-of-sight between communicants. Finally, the Nakagami- m model is a more generalist model that permits evaluating different environmental conditions by characterizing the effect of fading channels by tuning its m factor. Additionally, the basic Free Space model [186] is also considered. This model is characterized by the Friis formulation [186], which permits to study the signal propagation in an outdoor situation with neither obstacles nor other effects harming the transmitted signal (shadowing, fading, etc.).

As aforementioned, there are several propagation models that can be classified into different families. Examples of other well-known propagation models are: Weibull, Walfisch-Bertoni, COST-231, or Log-Normal shadowing models. For simplicity, in the following a brief review of the propagation models employed in this thesis is presented.

3.2.1 Free Space Model

The Free Space propagation model, also known as Friis model [186], is usually employed to predict the signal power level at the receiver side in a Free Space situation without obstacles between both extremes of the communication (see Fig. 3.1). It is a quite simple model since it only considers the distance as the main factor causing the signal level decreasing, ignoring other complex phenomena such as waves' reflection or diffraction processes.

The Free Space model is characterized by the Friis equation, which allows estimating the signal receiving power given the distance between transmitter and receiver, considering other basic parameter, too. The Friis equation is showed in expression (3.1),

$$P_r(d) = \frac{P_t G_r G_t}{L} \left(\frac{\lambda}{4\pi d} \right)^2 \quad (3.1)$$

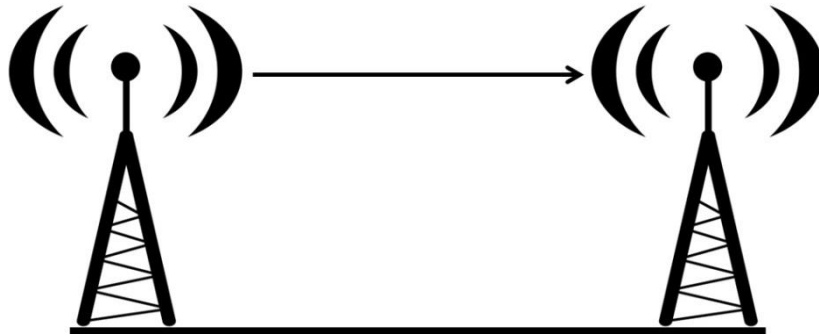


Figure 3.1. Free Space propagation environment.

where P_t is the transmitting power; G_r and G_t are the gain of the receiving and transmitting antennas, respectively; λ is the signal's wavelength and d is the distance between communicants, both of them in terms of meters. Finally L factor is not associated with the propagation process and represents the signal loss caused by transmission and reception electronic components (antennas, filters, transmission lines, etc.). The antenna's gain is a fundamental design parameter that is related to its physical dimensions, and it is calculated considering its effective area A_e , as shown in expression (3.2).

$$G = \frac{4\pi A_e}{\lambda^2} \quad (3.2)$$

It is important noting that the Friis formula for Free Space just holds for the far-field region (Fraunhofer region); it is in the far field that the propagating waves act as plane waves and the power decays inversely with distance. Hence, the Free Space propagation model is only applicable when the receiving antenna is located within the transmitting antenna's Fraunhofer region, whose boundary is determined by the expression (3.3)

$$d_f = \frac{2D^2}{\lambda} \quad (3.3)$$

where $d_f \gg \lambda$ and D is the greatest antenna's dimension. Therefore, the Free Space propagation model provides accurate predictions when the distance between communicants is greater than d_f .

In order to quantify the path-induced losses as a positive magnitude (in terms of dB), the expression (3.4) permit calculating the attenuation suffered by a wave in Free Space, given its wavelength and the covered distance.

$$P(dB) = -20 \log \frac{\lambda}{4\pi d} \quad (3.4)$$

Finally, Fig. 3.2 depicts the evolution of the signal's power level at the receiving side when the distance between both extremes of the communication is increased. It makes sense that as the distance separating transmitter and receiver increases, the received signal's level suffers greater decay.

3.2.2 Rayleigh Model

This probabilistic model characterizes one of the most common situations when evaluating mobile wireless networks, namely, when there is no line-of-sight between transmitter and receiver (see Fig. 3.3).

This model assumes that the signal's envelope amplitude has a Rayleigh probability density function (pdf) [183]. This happens because the received signal's envelope is composed by different multi-path distributions; it is well known that the envelope of the sum of two quadrature Gaussian noise signals obeys a Rayleigh distribution.

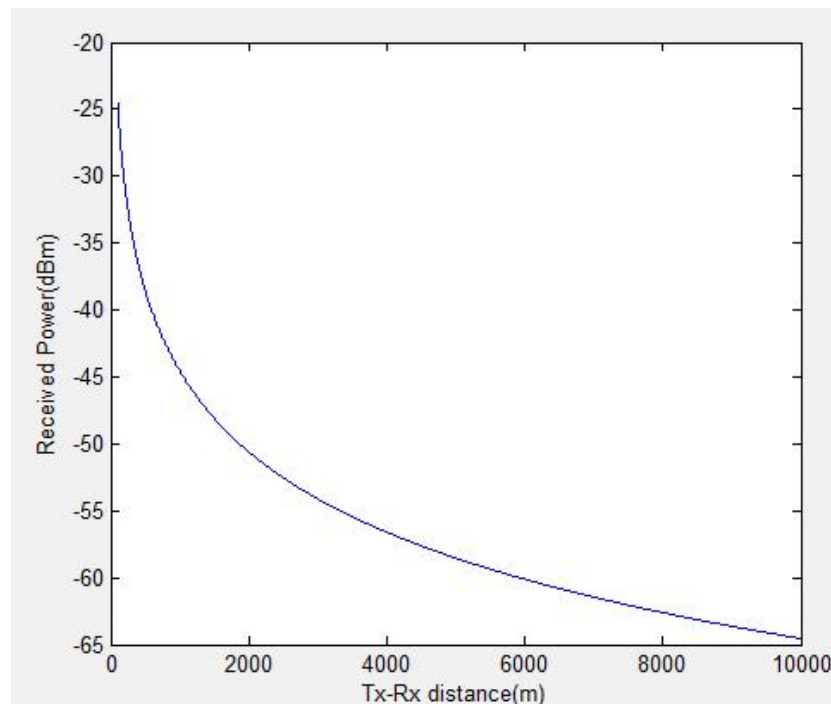


Figure 3.2. Received signal's level against distance between communicants. Free Space propagation model.

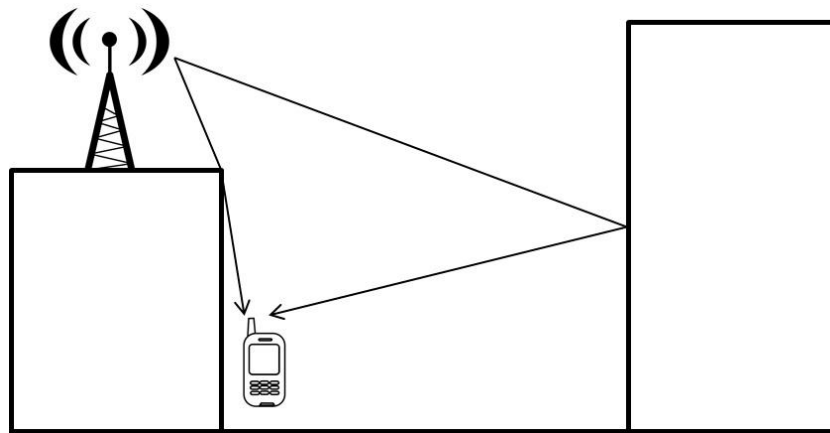


Figure 3.3. Propagation environment characterized by the Rayleigh model.

On one hand, the reception of signals formed by several multi-path components is the main basis to establish communication between devices that do not have line-of-sight, e.g., cellular networks. On the other hand, because of wave cancellation effects, these type of signals present variations in its envelope that difficult its proper reception and decoding. This phenomenon is referred to as short-term fading. The Rayleigh distribution has a pdf characterized by expression (3.5),

$$P(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(\frac{-r^2}{2\sigma^2}\right) & 0 \leq r \leq \infty \\ 0 & r < 0 \end{cases} \quad (3.5)$$

where r is the envelope amplitude of the received signal, and $2\sigma^2$ is the pre-detection mean power of the multi-path signal. Fig. 3.4 depicts the Rayleigh distribution's pdf as a function of σ .

The average value of this distribution is calculated by means of the expression (3.6).

$$r_{avg} = E[r] = \int_0^\infty r P(r) dr = \sigma \sqrt{\frac{\pi}{2}} = 1.2533\sigma \quad (3.6)$$

It is also interesting knowing the Rayleigh distribution's variance as it represents the AC power in the signal's envelope. Its calculation formula is showed in expression (3.7).

$$\sigma_r^2 = E[r^2] - E^2[r] = \int_0^\infty r^2 p(r) dr - \frac{\sigma^2 \pi}{2} = \sigma^2 \left(2 - \frac{\pi}{2}\right) = 0.4292\sigma^2 \quad (3.7)$$

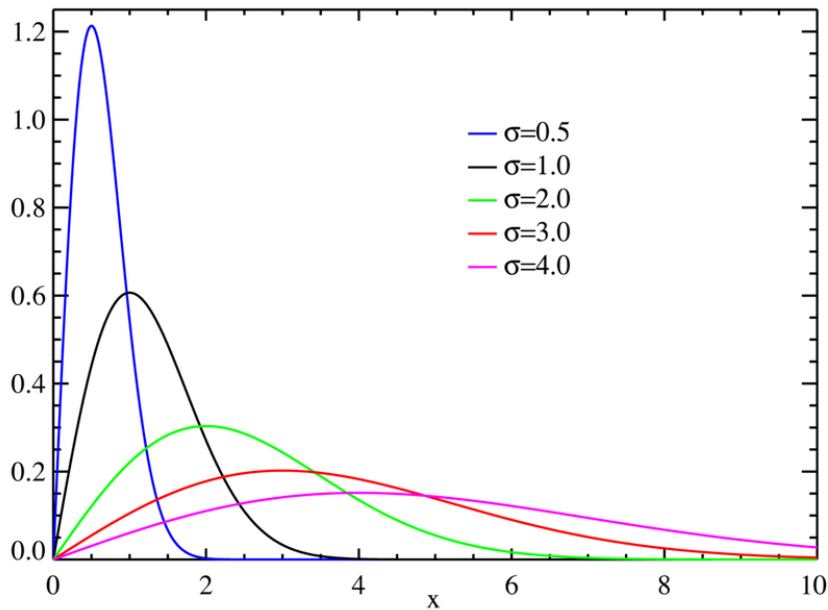


Figure 3.4. Rayleigh distribution's probability density function (extracted from [189]).

3.2.3 Rice Model

In realistic scenarios, it is often common that, in addition to multi-path contributions, there is a dominant stationary component in the received signal (specular component). This is usually caused by the existence of line-of-sight between transmitter and receiver (see Fig. 3.5). Thus, multi-path signals superpose over the specular component (nonfaded) and a new envelop is formed. This new envelop, affected by power variations caused by fading, has a Rician pdf [184]. As the amplitude of the specular component approaches zero, the Rician pdf degenerates to a Rayleigh pdf. Thus, as well as in the Rayleigh model case, signals characterized by the Rice model undergo increases and decreases in its power level due to the interference interactions among the different components of the wave.

The Rician pdf is determined by expression 3.8,

$$P(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(-\frac{r^2+A^2}{2\sigma^2}\right) I_0\left(\frac{Ar}{\sigma^2}\right) & A \geq 0, r \geq 0 \\ 0 & r < 0 \end{cases} \quad (3.8)$$

where r is the envelope's amplitude of the received signal, $2\sigma^2$ is the pre-detection mean power of the multi-path signal, A is the peak power of the specular component and I_0 is the modified Bessel function of the first kind with order zero. Fig. 3.6 shows the Rician distribution's pdf as a function of the dominant contribution's amplitude (A).

The Rician distribution can also be described in terms of a parameter, K , which is defined as the ratio between the deterministic signal power (A) and the variance of the multipath contributions ($2\sigma^2$), as shown in expression (3.9).

$$K = \frac{A}{2\sigma^2} \quad (3.9)$$

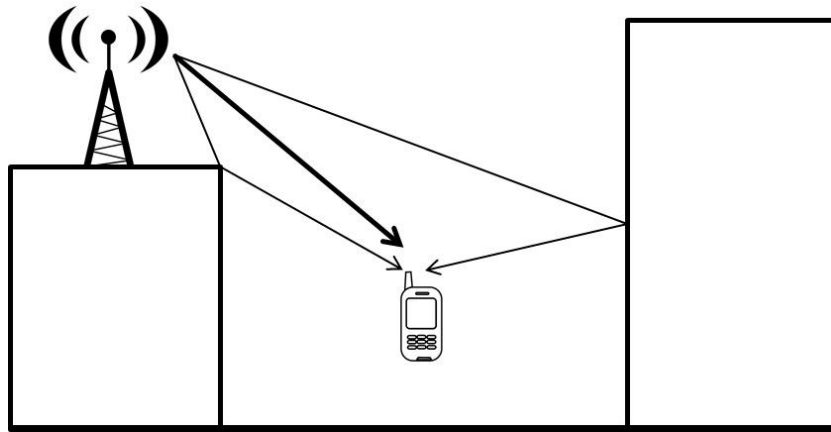


Figure 3.5. Propagation environment characterized by the Rice model.

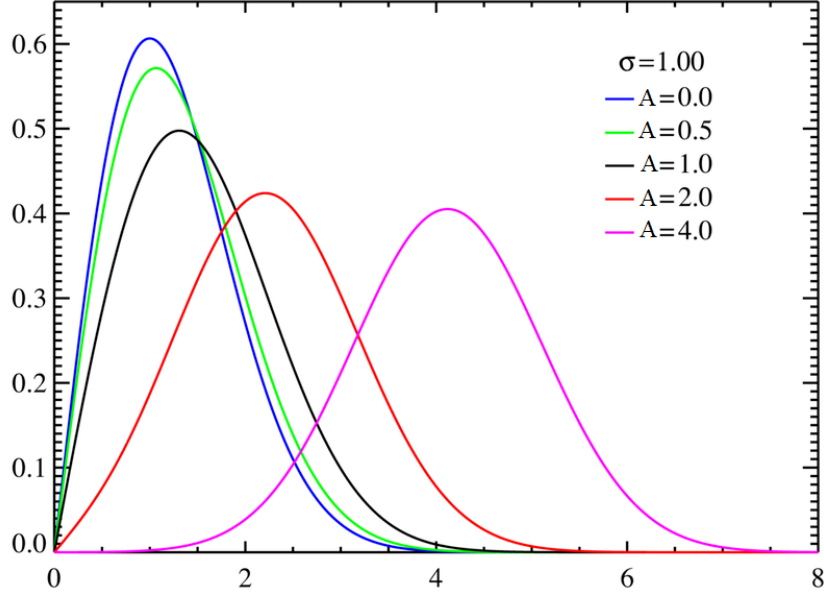


Figure 3.6. Rice distribution's probability density function (extracted from [189]).

The K factor, usually referred to as Rician factor, characterizes the Rician distribution. If the main contribution (specular component) decreases in amplitude ($A \rightarrow 0$), then the Rician factor (K) also tends to 0 and, as aforementioned, the Rician pdf approaches a Rayleigh pdf.

Finally, in order to provide a complete characterization of the Rice model, the expression (3.10) presents the received signal's envelop power pdf, $R(t)$, as a function of the K factor,

$$f_R(r) = \frac{2(K+1)r}{\Omega} \exp\left(-K - \frac{(K+1)r^2}{\Omega}\right) I_0\left(2r\sqrt{\frac{K(K+1)}{\Omega}}\right) \quad r \geq 0, K \geq 0, \Omega \geq 0 \quad (3.10)$$

where $\Omega = E[R^2]$ is the average signal power.

3.2.4 Nakagami- m Model

This model, which is based on the Nakagami- m distribution[185], is one of the most employed models to characterize the wireless physical layer. The main advantage that it provides is its generalization capability. As mentioned above, the Nakagami- m model is completely defined by the m factor, which characterizes the fading level of the environment. Thus, when $m = 1$, the Nakagami- m pdf is reduced to the Rayleigh pdf; with higher values of m the Nakagami- m pdf approaches the Rice pdf; finally, the fading level tends to 0 when $m \rightarrow \infty$. Therefore, the Nakagami- m model is suitable for describing statistics of mobile radio transmission in complex media such as the urban

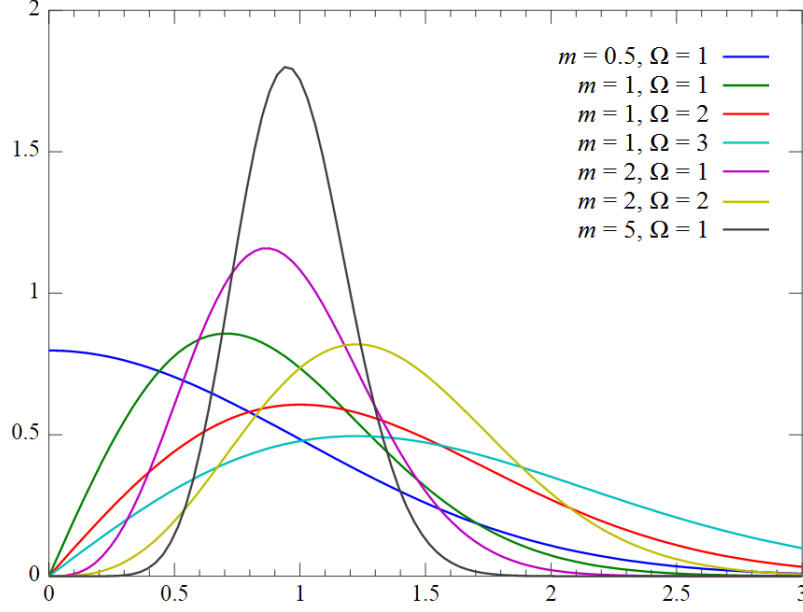


Figure 3.7. Nakagami- m distribution's probability density function (extracted from [189]).

environment or indoor situations. Since the Nakagami- m random process is defined as an envelope of the sum of $2m$ independent Gauss random processes, the Nakagami- m distribution is described by the pdf (3.11),

$$p_z(z; m, \Omega) = \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m z^{2m-1} \exp\left(-\frac{m}{\Omega} z^2\right), z > 0, m \geq \frac{1}{2} \quad (3.11)$$

where z is the received signal level, $\Gamma(\cdot)$ is the gamma function, and represents the average signal power ($\Omega = E[R^2]$). Observe that the Nakagami- m distribution is also a function of the m factor, which is formally defined as showed in expression (3.12).

$$m = \frac{E^2[z]}{Var[z^2]} \quad (3.12)$$

Fig. 3.7 shows the Nakagami- m distribution's pdf as a function of m and Ω .

3.3 VoIP QoE Evaluation in 802.11g Infrastructure Networks

Currently, the most widespread wireless standard to configure Wireless Local Area Networks (WLANs) is the IEEE 802.11 family (Wi-Fi). Specifically, Wi-Fi networks configured in infrastructure mode are the most employed systems to provide Internet access to end-users in both public and private domains. This operational mode consists

of a central node, i.e., an access point, that provides network services to the rest of the network nodes, e.g., routing, IP configuration, etc. The remaining nodes act as hosts, sending and receiving data traffic through the central access point. Given the expansion of these wireless networks, it is convenient studying their capability to support highly demanded services by the end-users, such as VoIP. Some relevant examples of works that evaluate the capacity for VoIP traffic in wireless local area networks by simulation or experimentation are [190]–[196]. In these works, the effect of layer two handoffs, delay constraints, channels with constant bit error rates, packet size, UDP traffic, TCP traffic, overheads from different layers, etc., were intensively evaluated. From their results, it is observed that the capacity achieved (defined as number of valid VoIP calls) is different, of course, depending on the simulated and/or configured network conditions (testbed). It is worthy to mention the work done by Shin and Schulzrinne [190], which studied the capacity of VoIP traffic in an IEEE 802.11b network for a wide range of network parameters. They showed that there are several factors such as the buffer size at the access point, the type of retry limit of the IEEE 802.11, or the preamble size that have a significant impact on the performance of VoIP, in particular, limiting the number of VoIP calls supported by the network. By using the G.711 codec [197] and a packetization interval of 20 ms, they achieved a capacity of 15 constant bit-rate calls and 38 calls in variable bit-rate VoIP traffic; understanding as maximum capacity the number of VoIP calls so that the 90th percentile of the one-way end-to-end delay is less than 150 ms. However, it should be noted that neither this work nor the previous ones detailed the effect of more realistic physical layers. The physical layer plays a key role that cannot be ignored when evaluating QoE, since packet losses are directly impacted by this factor ignoring its origin, and QoE models usually include packet losses as one of their input parameters.

To the best of the author's knowledge, only the works in [198] and [199] tackled the physical layer effect. In [198], the quality of voice communication over IEEE 802.11a networks is studied. Particularly, their authors included in the simulations a simplified version of the Keenan-Motley partition loss model incorporating a Free Space propagation model and shadowing, representing an office scenario with no line-of-sight. As a result, they showed the cumulative distribution functions of packet error rate for different Signal-to-Noise ratios (distances) and different data rates. Using stationary VoIP users subject to a particular fading realization, they estimated QoE values (in

terms of MOS) for the VoIP calls using the E-model. They showed how the QoE is much greater if packet errors from fading are not taking into account. A more complete work was done in [199], where authors examined the behavior of 802.11s wireless mesh networks including Nakagami- m fading channels. They showed how the SNR should be increased to achieve similar QoE levels in VoIP calls than those achieved with Additive Gaussian Noise channels. In their study, they used the E-Model for QoE evaluation.

In contrast to these works, in the following sections it is evaluated the effect of fading on the QoE of VoIP calls, including a more realistic and complete physical layer characterization represented by the use of four propagation models. It is compared delay, packet loss, and QoE (in terms of MOS obtained with the E-model [39]) to assess the maximum allowed capacity, i.e., not exceeding the recommended limits for delay and packet losses according to the ITU-T Rec. G.114 [182] and G.1010 [16], and reaching a QoE value that represents an “acceptable for most users” quality level.

The following evaluation of Wi-Fi infrastructure networks supporting VoIP traffic is separated in two parts. First, the network performance with two different codecs (G.711 A-law at 64 Kbps and G.726 at 24, 32, or 40 Kbps), two coding packetization-intervals (10 ms and 20 ms), and also two different available wireless bandwidths (11 Mbps and 54 Mbps) is evaluated, considering three wireless channel conditions (characterized by the Rayleigh, Rice, and Free Space models). The second part is specifically focused on assessing the impact of fading on the system performance. By means of the Nakagami- m model, its shape factor, m , is tuned, thus characterizing the wireless transmission medium with different levels of “hostility”, in terms of fading. Hence, by considering similar scenarios to those evaluated in the first part, the system performance is studied, regarding the impact of fading channels on the wave propagation.

Prior to the mentioned evaluation, an analytical characterization of the theoretical capacity of an 802.11g infrastructure network, in terms of VoIP calls, is developed. Additionally, a brief description of the QoE estimation model used to evaluate the quality of the VoIP calls is also presented. Finally, the employed simulation test-bench is explored too.

3.3.1 System Theoretical Capacity

Following the methodology proposed by Shin and Schulzrinne [190], the maximum capacity (C), in terms of simultaneous VoIP calls, that the system under evaluation can

support is calculated. This task is performed by means of the expression (3.13). In this formula, P is the packetization time-interval (s), T_{DIFS} (s) and T_{SIFS} (s) represent the 802.11 Distributed Interframe Space (DIFS) and Short Interframe Space (SIFS), respectively; T_V (s) is the necessary time to transmit a VoIP packet, including all headers, i.e., PLCP, MAC, IP, UDP, and RTP. Finally, T_{ACK} (s) and T_{BO} (s) are the time employed to send an 802.11 ACK frame and the back-off period, respectively. T_{BO} is calculated as the number of back-off slots times T_s , where T_s (s) is a slot-time, and the number of back-off slots has a uniform distribution over $(0, CW_{min})$, with an average of $CW_{min}/2$. CW_{min} is the 802.11 minimum content window. Note that for this calculation, constant transmission rate is assumed, i.e., no silence detection algorithms are employed during the voice coding process.

$$C = \frac{P}{T_{DIFS} + T_{SIFS} + T_V + T_{ACK} + T_{BO}} \quad (3.13)$$

By setting all the 802.11 parameters to those shown in Table 3.3, the system theoretical capacity is shown in Table 3.1. Although the similar capacity figures attained for the different coding bit-rates may be surprising, these values are justified by (i) the impact of the headers added during the encapsulation process and (ii) the waiting time-intervals related to the medium access scheme. Time-intervals such as SIFS, DIFS or T_{BO} present great values that mask the effect of employing low bit-rate voice codecs.

3.3.2 E-Model Implementation

One of the objectives of this chapter is evaluating the effect of the impairments introduced by the wireless propagation on the QoE of VoIP communications. To this

TABLE 3.1. 802.11 infrastructure network theoretical capacity (in terms VoIP calls)

Codec	Tx Rate			
	11 Mbps		54 Mbps	
	Packetization time-interval		Packetization time-interval	
	10 ms	20 ms	10 ms	20 ms
G.711	14	26	24	46
G.726 (40 kbps)	15	28	24	47
G.726 (32 kbps)	15	29	24	48
G.726 (24 kbps)	15	29	24	48

end, it is necessary selecting the most proper methodology to estimate the quality of this service from an end-user perspective. For that reason, in the previous chapter an extended review of the different models developed to estimate the QoE of VoIP real-time transmissions was presented. Thereby, the chosen methodology to evaluate the quality of the calls is the reduced version of the E-Model [39]. This model is elected due to (i) its computational simplicity that permits QoE estimation in real-time, and (ii) the wide bibliography related to it, which ensures accurate results in the scenario under evaluation. The reduced E-model algorithm was implemented as a standalone Omnet++ module that could be located at any point of the network. Fig. 3.8 shows a descriptive flow-chart of the E-model implementation embedded in the simulator. This design permits evaluating the VoIP traffic in different transit zones, allowing also to find failure points or bottle-necks. Given the employed topology (see Fig. 3.9) the QoE estimator was finally located in each of the VoIP receivers (RX). The calculation of the factor R was performed by using expression (3.14),

$$R = R_o - I_{dd} - I_{e-eff} \quad (3.14)$$

where R_o takes the value 93.4. In turn, I_{dd} is obtained by applying the expression (3.15),

$$I_{dd} = \begin{cases} 0 & T_a < 100 \text{ ms} \\ 25 \left[(1 + X)^{\frac{1}{6}} - 3 \left(1 + \left(\frac{X}{3} \right)^6 \right)^{\frac{1}{6}} + 2 \right] & T_a \geq 100 \text{ ms} \end{cases} \quad (3.15)$$

with:

$$X = \frac{\log\left(\frac{T_a}{100}\right)}{\log 2} \quad (3.16)$$

where T_a is the connection's one-way delay. Finally, I_{e-eff} is attained by using the expression (3.17),

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}} \quad (3.17)$$

where P_{pl} is the probability of packet loss and $BurstR$ is the burst ratio, which is used to capture the burst conditions of the packet loss distribution. In this case it takes a value of 1, as random losses are assumed. As discussed in the previous chapter, the values of I_e and B_{pl} depend on the employed coding scheme. Taking into account the considered codecs (G.711 and G.726), the applied values for these parameters are shown in Table 3.2 (extracted from the ITU-T Rec. G.113 [42]).

TABLE 3.2. Values for parameters I_e and B_{pl}

Codec	I_e	B_{pl}
G.711	0	4.3
G.726 (40 kbps)	7	24
G.726 (32 kbps)	12	24
G.726 (24 kbps)	25	38

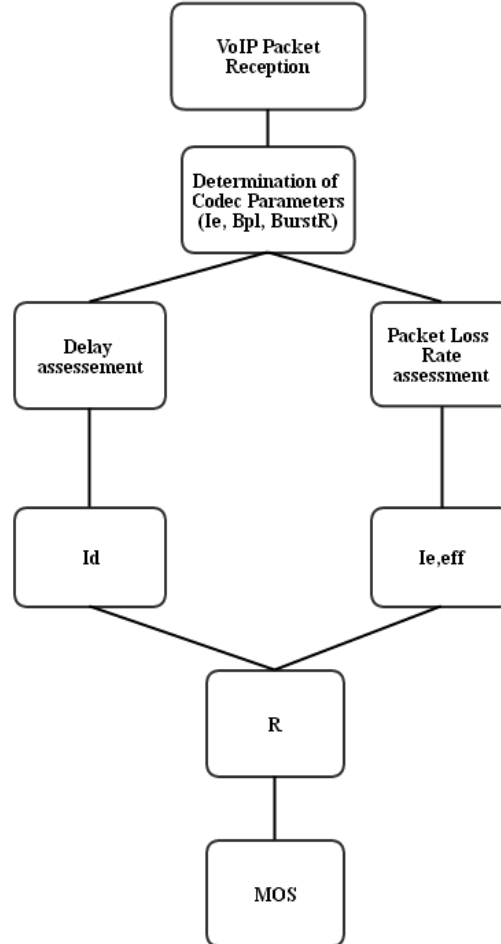


Figure 3.8. E-model flow-chart.

3.3.3 Simulation Environment

The simulation study was done with Omnet++ v4.2.2 and the InetManet framework [10]. As part of the InetManet framework, the VoIP traffic generator creates a stream of VoIP packets from an arbitrary sound file; as input, a set of the ITU-T Rec. P.50 test signals was used. These audio files include a voice pattern with voice activity and silences. This feature, in addition to the silence detector employed by the VoIP transmitters, permitted the generation of a Variable Bit Rate (VBR) flow from transmitter (TX) to receiver (RX). In the simulations, it was used the Pulse Code

Modulation (PCM) [200] and the Adaptive Differential Pulse Code Modulation (ADPCM) [201], currently defined as the ITU-T standards (i) G.711 A-law (64 Kbps) and (ii) G.726, respectively. The latter employs three different coding rates (24, 32, and 40 Kbps). In order to assess the quality of the VoIP calls, and considering different methodologies found in the literature [15, 16], finally, an E-Model implementation was added to the VoIP receivers following the guidelines of ITU-T Rec. G.107 and G.113 (see next section).

The simulated scenario was a wireless IEEE 802.11g infrastructure network. Two different transmission rates included in the 802.11g standard (11 and 54 Mbps) were used to study the performance for different available bandwidths. The Ethernet-to-wireless network topology illustrated in Fig. 3.9 was employed. The access point buffer was set to store 100 frames and the retry limit was established as the 802.11 standard short value, i.e., 7. The wireless card modules were set to a transmission power of 5 mW and a sensitivity of -85 dBm, with a carrier frequency given by the 802.11 standard, i.e., 2.4 GHz, and a signal-to-noise threshold of 4 dB. These are the usual values recommended, for instance by the Cisco Aironet Series wireless adapters among other manufacturers. The effect of fading channels were assessed by carrying out simulations using the Rice model to simulate an indoor situation with a line-of-sight between transmitter and receiver, and the Rayleigh model aiming at simulating an indoor situation without a line-of-sight between transmitter and receiver. In order to provide deeper insights about the impact of fading channels on the wireless system performance, the Nakagami- m propagation model was also employed. As discussed above, this model is characterized by the shaping-factor m : the smaller the m value, the greater the level of fading. If $m = 1$, the Nakagami- m model represents the Rayleigh model. When $m > 1$, the Nakagami- m closely approximates to the Rice model, which introduces less fading in the transmission channel. All these results, which take into account the impact of the wireless transmission channel, are compared with those attained in scenarios characterized by the Free Space propagation model, in which the physical layer has minimum influence on the wave propagation. Taking into account the described radio parameters and the four propagation models under study, the system was evaluated for a variable distance between the wireless VoIP nodes (TX in Fig. 3.8) and the access point, obtaining a network coverage ranging from 25 m to 475 m, depending on the propagation model and the bandwidth employed. It was assumed a

TABLE 3.3. 802.11g parameters

Parameters	Bytes	Time	
		11 Mbps	54 Mbps
SIFS, DIFS, SLOT (μ s)	-	{ 10, 50,20}	{ 10, 50, 20}
CW _{MIN} (slots)	-	31	31
PLCP preamble (μ s)	-	144	4
{PLCP, MAC, SNAP} headers (μ s)	-, 28, 8	{48, 20.36, 5.81}	{16, 4.15, 1.18}
IP + UDP + RTP headers (μ s)	40	29.09	5.92
Voice (G.711, 10 ms) (μ s)	160	58.18	11.85
Voice (G.711, 20 ms) (μ s)	160	116.36	23.70
Voice (G.726-40, 10 ms) (μ s)	100	36.36	7.4
Voice (G.726-40, 20 ms) (μ s)	100	72.72	14.8
Voice (G.726-32, 10 ms) (μ s)	80	29.09	5.92
Voice (G.726-32, 20 ms) (μ s)	80	58.18	11.84
Voice (G.726-24, 10 ms) (μ s)	60	21.81	4.44
Voice (G.726-24, 20 ms) (μ s)	60	43.62	8.88
ACK (μ s)	14	10.18	2.07

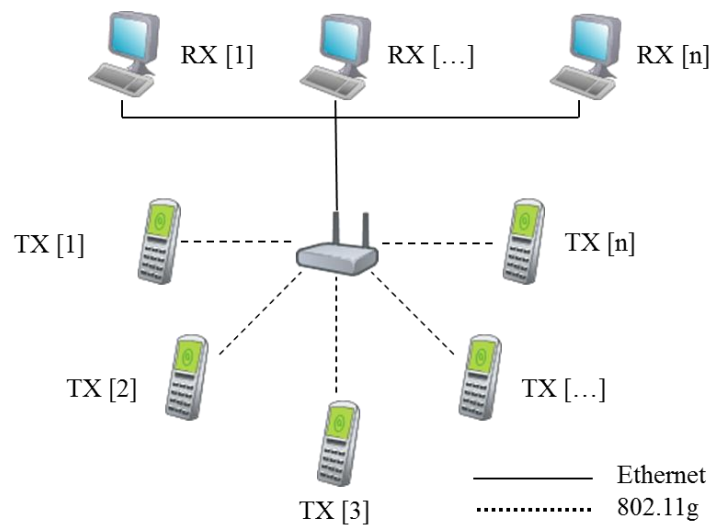


Figure 3.9. Network under evaluation.

VoIP header, i.e., RTP (Real Time Protocol), of 12 bytes. In the simulations, two different packetization time lengths, namely, 10 ms and 20 ms, were employed. The starting time for each VoIP call was chosen randomly with a uniform distribution in a time range of (0, 8) s. The audio sources lasted 8 s. Additional 802.11g parameters were set according to Table 3.3.

3.3.4 Results

In this section, the simulation results obtained to study the VoIP system behavior for different scenarios are shown. The metrics under consideration are the QoE in terms of MOS, the delay, and the Packet Loss Ratio (PLR) for every call, aiming at analyzing how the distance, and therefore the fading channels, affects the system, in terms of capacity and conversational quality. In order to set each individual call as valid, three metrics are usually used in the bibliography: QoE in terms of MOS, one-way delay, and/or PLR. Following the guidelines of the ITU-T Rec. G.114 [182] and G.1010 [16], in this study a call is defined as valid if the final QoE (MOS) value attained for this call is over 3.1 (see Table 2.4). Furthermore, the maximum one-way delay introduced by the network and PLR accepted are 80 ms and 5%, respectively. The former is calculated as follows; the one-way end-to-end delay of voice packets is supposed to be less than 150 ms (ITU-T Rec. G.114). Assuming the codec delay to be about 30-40 ms at both sender and receiver, and that the Ethernet connections are simulated as ideal links, i.e., do not add extra delay to the communication, then the wireless network should contribute less than 80 ms delay.

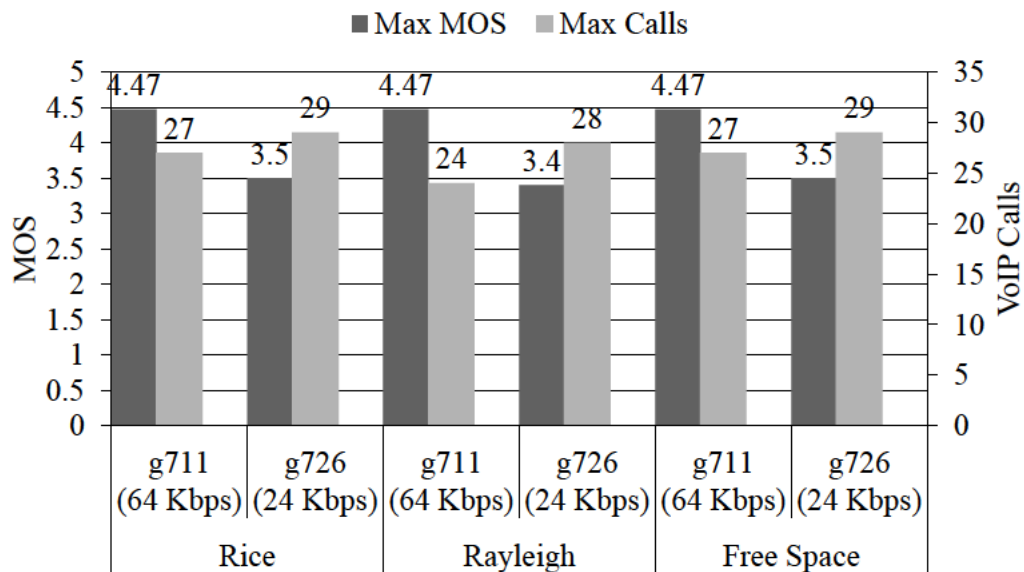
As pointed above, the network evaluation has been divided in two parts. First, the system performance with different voice-coding configurations and two different available wireless bandwidths (11 Mbps and 54 Mbps) is studied. In order to provide the most realistic results, three wireless propagation models are considered, namely, Rayleigh, Rice, and Free Space models. The second part is strictly focused on measuring the impact of fading channels on the system performance. Thus, the same scenarios studied in the first part are considered, but, by means of the Nakagami- m model, the physical layer's hostility is tuned in order to understand the impact of fading on the VoIP service quality.

3.3.4.1 Regular Network Evaluation: Rice and Rayleigh models

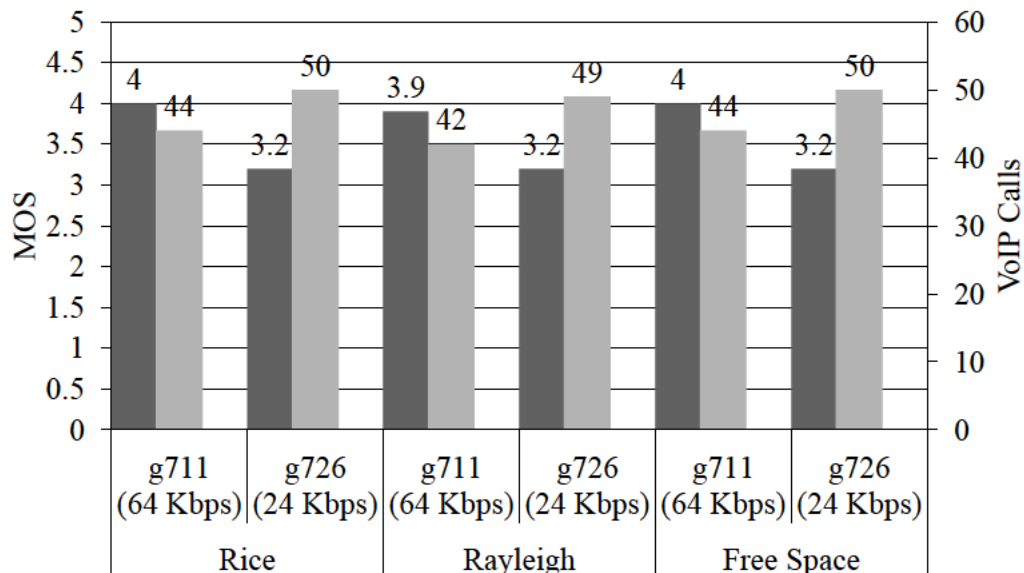
As aforementioned, two different VoIP packetization time-intervals (10 ms and 20 ms) have been employed to study the response of the IEEE 802.11 infrastructure network under consideration. Thus, this section is sub-divided, aiming to show separately the obtained results by employing each of the mentioned packetization intervals.

3.3.4.1.1 Packetization of 20 ms

Regarding the effect of the codec on the maximum VoIP system capacity, it is noticed that more calls are valid for G.726 (a low bit-rate codec) than for G.711. Fig. 3.10 shows a comparison between the maximum number of valid VoIP calls and the average QoE (MOS) value attained, for the three propagation models, the two transmission rates used, and the two codecs G.711 (64 Kbps) and G.726 (24 Kbps) under consideration. Recall that a valid VoIP call is interpreted as achieving a QoE



(a) 11 Mbps



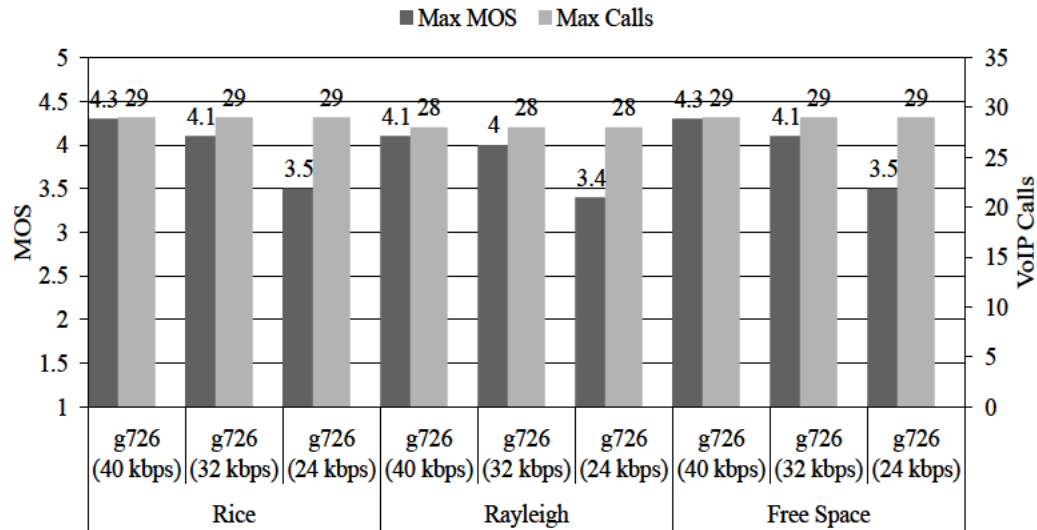
(b) 54 Mbps

Figure 3.10. Comparison of maximum VoIP capacity and QoE (MOS) obtained in the VoIP calls and codecs G.711 A-law and G.726 (24 Kbps) for the different propagation models under study.

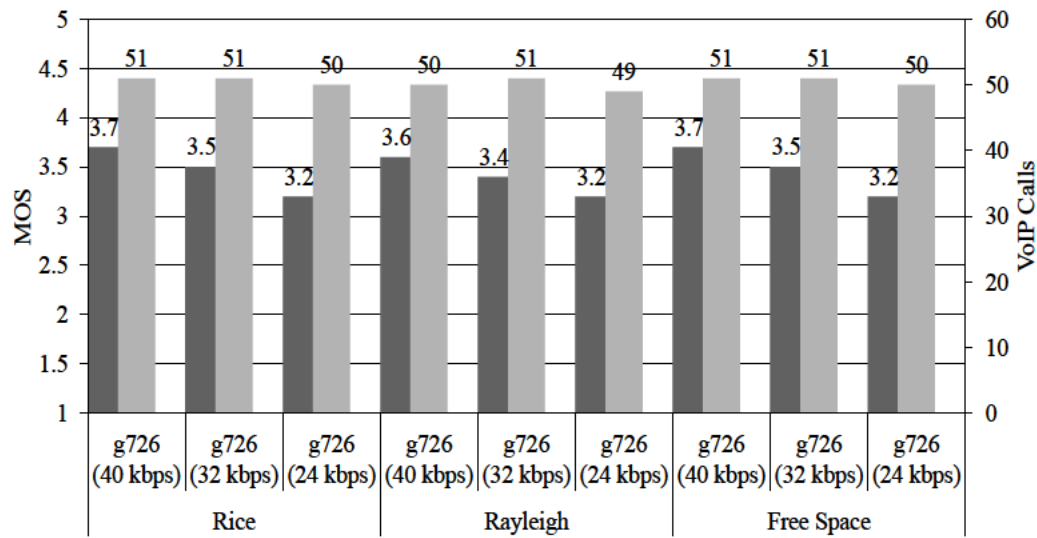
(MOS) equal or greater than 3.1. At 11 Mbps (Fig. 3.10 (a)), 27 valid calls are achieved for G.711 and 29 calls for G.726 (24 Kbps) for both Rice and Free Space scenarios. However, only 24 calls for G.711 and 28 calls for G.726 (24 Kbps) are achieved in a Rayleigh environment. In a Free Space scenario at 54 Mbps (see Fig. 3.10 (b)), G.711 achieves 44 valid calls against 50 calls for G.726 (24 Kbps), whereas only 42 and 49 calls are accepted as valid for each codec assuming the Rayleigh model. It is clear that the low bit-rate codec (G.726) supports more valid VoIP calls.

However, observe that the average QoE (MOS) reached with G.726 is noticeably smaller than the QoE (MOS) obtained in VoIP calls with G.711. That is, there exists a trade-off between the maximum number of valid VoIP calls supported by the wireless system and the QoE level provided in those calls: the more number of calls the smaller the QoE. This same behavior is perceived in Fig. 3.11, where the capacity and QoE (MOS) for the three different coding rates of the G.726 codec are represented. It can be seen that the capacity does not exhibit a significant increase regarding coding rates. Indeed, for 11 Mbps (Fig. 3.11 (a)) the capacity is the same for all coding rates. It is worthy to mention the results obtained for 54 Mbps (Fig. 3.11 (b)). Note how the capacity reached using a coding rate of 24 Kbps is lower than that obtained with the other coding rates. This is due to the low QoE (MOS) value attained for VoIP calls in this scenario, which makes us to discard active VoIP calls because their level of quality are lower than the QoE (MOS) limit established, i.e., 3.1.

On the other hand, Fig. 3.12 and Fig. 3.13 show the one-way delay and PLR, respectively, for the VoIP calls in order to study the effect of fading channels over the key network metrics affecting the communication quality. Observe that Fig. 3.12 (a) and Fig. 3.12 (b) depict the evolution of QoE (MOS) and delay as the number of calls accessing the system increases for Rice and Rayleigh propagation models, respectively. These metrics are represented for the transition distance for each scenario, i.e., when the system capacity decays from its maximum to lower levels. As mentioned previously, the maximum one-way delay limit was set to 80 ms. It is observed that when the delay exceeds that value the QoE (MOS) indeed drops under 3.1, so those VoIP calls are not accepted as valid ones. The same behavior is detected by analyzing the PLR. It was set a PLR limit of 5%, meaning that communications suffering PLR over this limit should not be accepted as valid. As shown in Fig. 3.13 (a) and Fig. 3.13 (b), when the PLR exceeds that threshold, quality estimation fell below 3.1, so these calls are also set as



(a) 11 Mbps

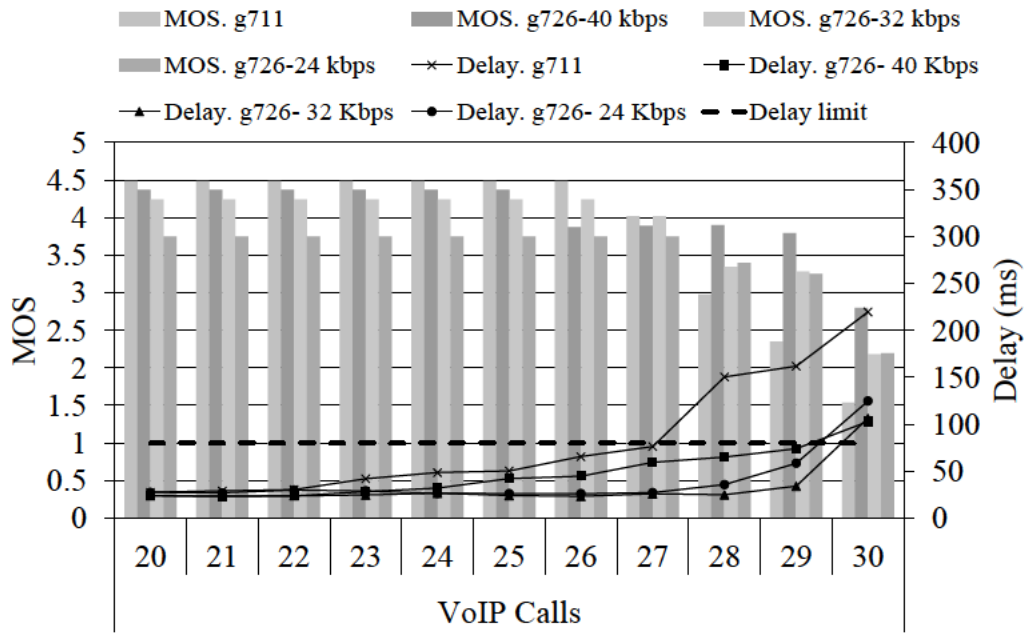


(b) 54 Mbps

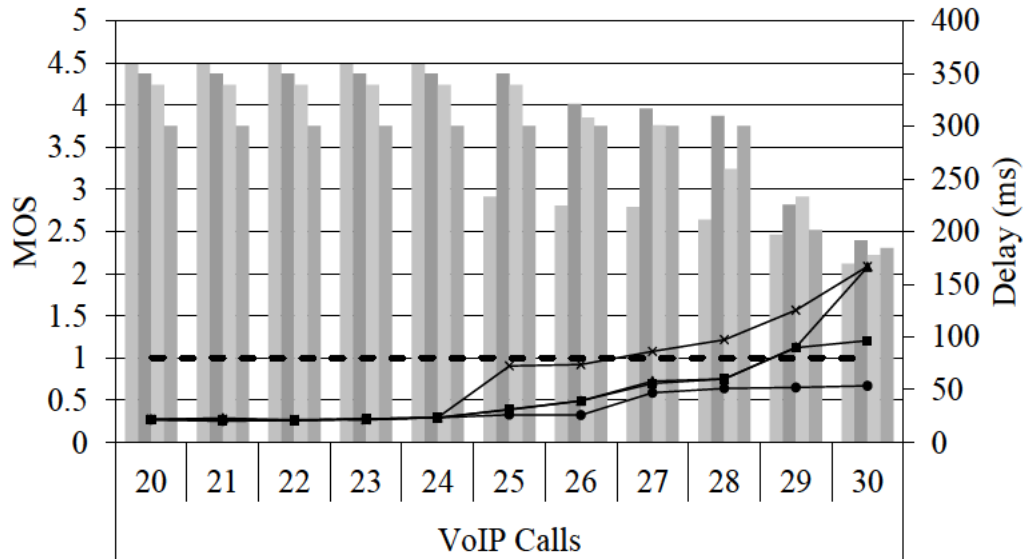
Figure 3.11. Comparison of maximum VoIP capacity and QoE (MOS) obtained in the VoIP calls and codec G.726 at 24, 32, and 40 Kbps for the different scenarios under study.

unacceptable. In consequence, the assumption that a QoE (MOS) value of 3.1 is accurate to establish the maximum number of valid VoIP calls accepted by the system is correct, even though this value already means that many users are dissatisfied with the service (i.e., it is not very strict with the quality level demanded). Note that the QoE (MOS) value 3.1 means that either delay or PLR are above the recommended limits, but it does not mean that both parameters are simultaneously exceeding their thresholds.

Regarding the effect of fading channels over the system coverage range and capacity, Fig. 3.14 and 3.15 shows how the number of simultaneous valid VoIP calls supported by the network decreases as the distance between the wireless VoIP sources and the



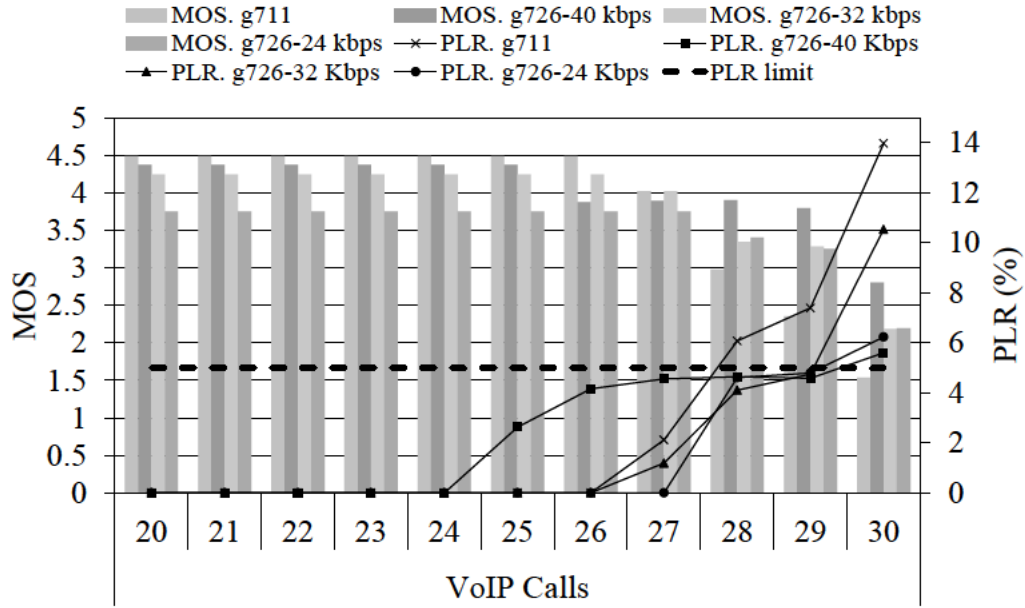
(a) Rice model



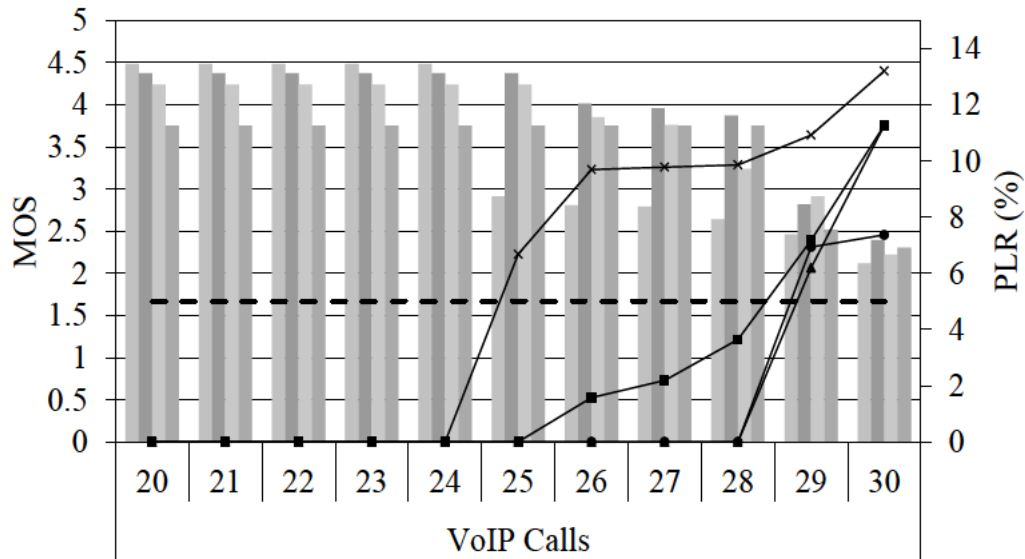
(b) Rayleigh model

Figure 3.12. QoE (MOS) and delay for 11 Mbps and packetization of 20 ms. Note that for each group of bars, the bar sequence is always as follows: G.711 (64 Kbps), G.726 (40 Kbps), G.726 (32 Kbps), and G.726 (24 Kbps).

access point increases. Different system responses are noticed depending on the propagation model under consideration. First, the Free Space model reaches the longest distance with valid calls (375 m) disregarding the codec in use and the network bandwidth (observe that there is no difference in the coverage range of G.711 or G.726 at 24 Kbps in both Fig. 3.14 and Fig. 3.15). Second, probabilistic models show a working range much lower than that obtained for the Free Space model, as it could be



(a) Rice model

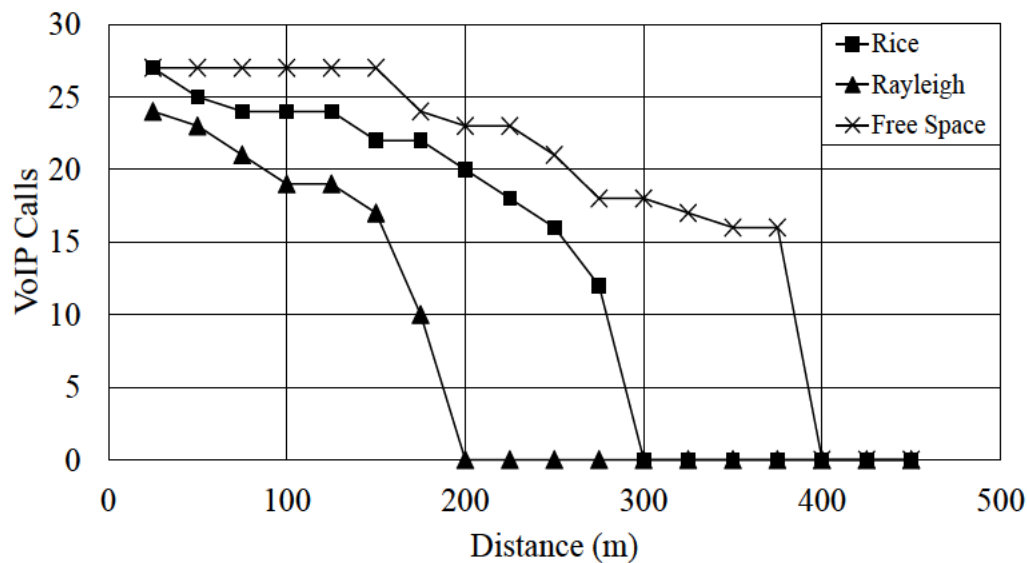


(b) Rayleigh model

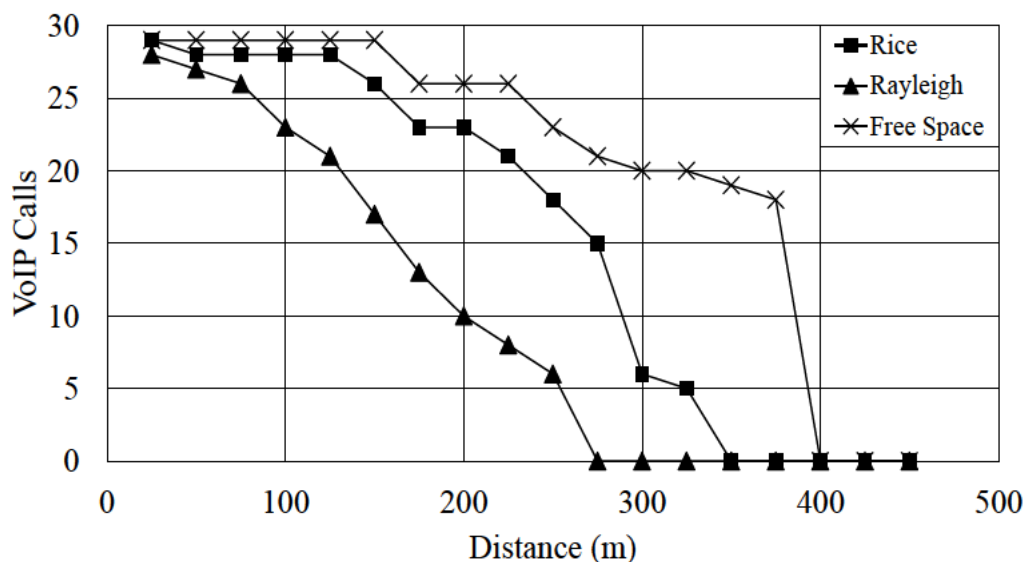
Figure 3.13. QoE (MOS) and PLR for 11 Mbps and packetization of 20 ms. Note that for each group of bars, the bar sequence is always as follows: G.711 (64 Kbps), G.726 (40 Kbps), G.726 (32 Kbps), and G.726 (24 Kbps).

expected; for instance, 275 m and 175 m for Rice and Rayleigh models respectively at 11 Mbps and codec G.711 (see Fig. 3.14 (a)). Longer distance ranges with valid calls are accomplished at 54 Mbps (Fig. 3.15 (a) and Fig. 3.15 (b)), with 300 m and 200 m for Rice and Rayleigh scenarios respectively disregarding the codec in use. The existing decrease on the work range for the Rice and Rayleigh physical layer scenarios is explained by the impact of the fading channels introduced by these models. Rayleigh

environments show a more pronounced reduction on the capacity as the distance increases. This model assumes the inexistence of dominant contributions along a line-of-sight, so that, the contributions that arrived to the wireless nodes are weaker than for the other models. In other words, Rayleigh is the model that introduces more fading in the transmission channel. Meanwhile, Rice scenarios support valid VoIP calls on greater ranges than Rayleigh's, although still far from the distances reached with Free Space model. It can be also seen from the results in Fig. 3.14 and 3.15 that low bit-rate codecs not only allow a higher maximum number of valid VoIP calls, but the working range



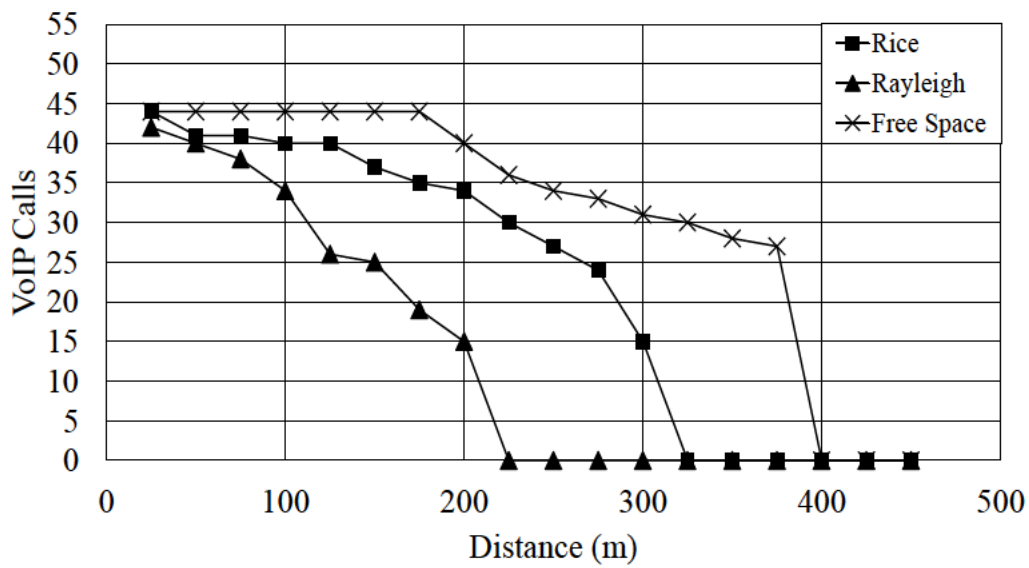
(a) Codec G.711



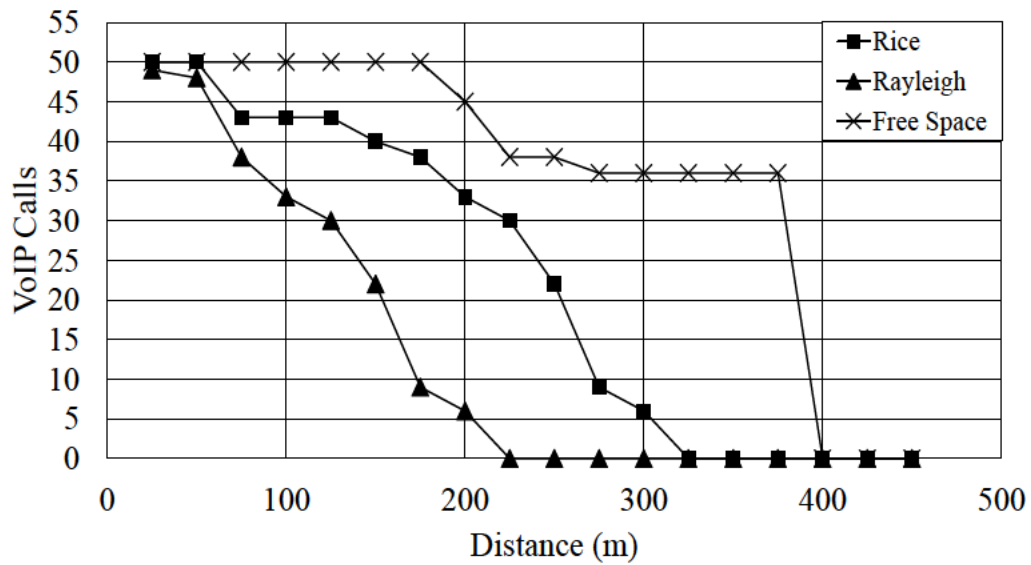
(b) Codec G.726 (24 Kbps)

Figure 3.14. VoIP calls supported by the IEEE 802.11g wireless network at 11 Mbps and packetization of 20 ms.

with valid VoIP calls is larger than the one obtained with G.711 under the same propagation model. It is also observed that for the G.711 codec, the fact of being in a wireless network with more bandwidth slightly affects the working range, presenting an improvement of 25 m between 11 Mbps and 54 Mbps (see Fig. 3.14 (a) and Fig. 3.15 (a)). In contrast, for the G.726 codec at 24 Kbps, the more the bandwidth available the smaller the working range. For instance, G.726 with Rayleigh and 11 Mbps present a working range of 250 m (Fig. 3.14 (b)), whereas the working range is only 200 m at 54 Mbps (Fig. 3.15 (b)). The reason can be found in the low level of QoE (MOS) achieved



(a) Codec G.711



(b) Codec G.726 (24 Kbps)

Figure 3.15. VoIP calls supported by the IEEE 802.11g wireless network at 54 Mbps and packetization of 20 ms.

by this codec. Summarizing, fading affects in two ways the QoE of the VoIP transmission independently of the codec in use: the more the fading (i) the smaller the maximum distance reached with valid calls and (ii) the smaller the number of valid VoIP calls for the same distance.

3.3.4.1.2 Packetization of 10 ms

A packetization time-interval of 10 ms was also employed with the aim of studying the performance of the network supporting smaller VoIP packets. Focusing on the quality of the VoIP communication, as shown in Table 3.4, similar QoE (MOS) values are obtained as in the greater packetization time-length for the most hostile scenario, i.e., that characterized by the Rayleigh propagation model. The behavior of the QoE is the same as for 20 ms: the higher environmental hostility, the greater drop in the calls quality; it is also obtained a generalized drop in the QoE in the 54 Mbps scenarios compared with those at 11 Mbps.

Additionally, the system coverage range is neither affected by the size of packets, obtaining similar work ranges in both, fading and non-fading scenarios, being the latter those with the greatest coverage ranges. On the other hand, the system capacity suffers a deep drop by employing the smallest packetization time-interval. This happens due to the decrease on the length of the voice payload, which provoked that the legacy preambles and overheads gain weight in front of the VoIP data; consequently, the wireless transmission channel is inefficiently utilized. Table 3.5 shows the system capacity obtained in the Rayleigh scenario using both packetization time-durations. Observe how the difference in the system capacity, by employing the no-compression codec or the low-bit rate codec, decreases using the packetization time-length of 10 ms.

TABLE 3.4. Maximum quality of the VoIP calls (MOS) for the Rayleigh model obtained through simulation

Codec	Tx Rate & Packetization interval			
	11 Mbps		54 Mbps	
	10 ms	20 ms	10 ms	20 ms
G.711	4.48	4.47	4.48	3.9
G.726 (40 Kbps)	4	4.1	3.8	3.6
G.726 (32 Kbps)	3.7	4	3.5	3.4
G.726 (24 Kbps)	3.3	3.4	3.4	3.2

TABLE 3.5. Maximum capacity of the system for the Rayleigh model obtained through simulation (in VoIP calls)

Codec	Tx Rate & Packetization interval			
	11 Mbps		54 Mbps	
	10 ms	20 ms	10 ms	20 ms
G.711	13	24	21	42
G.726 (40 Kbps)	15	28	25	50
G.726 (32 Kbps)	15	28	25	51
G.726 (24 Kbps)	16	28	24	49

As discussed in the theoretical characterization of the network capacity, a small size of the VoIP payload masks the capacity-gain obtained by the low-bit rate codec using the packetization interval of 20 ms.

3.3.4.2 Fading Impact Evaluation: Nakagami- m model

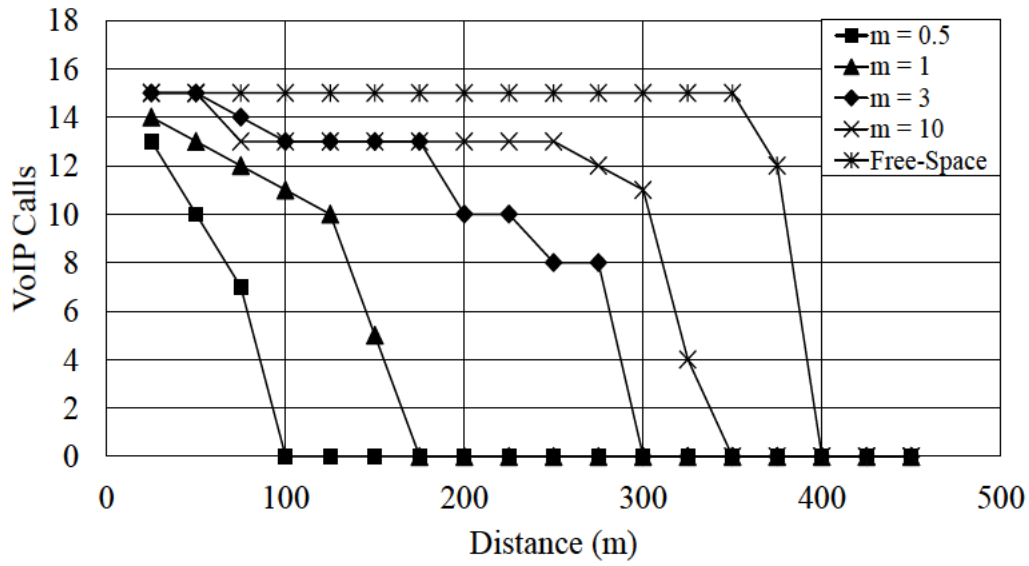
In the following, the simulation results obtained to study the VoIP system behavior under different fading conditions by using the Nakagami- m propagation model are shown. As in the previous section, the outcomes are separately analyzed depending on the two considered packetization intervals (10 ms and 20 ms).

3.3.4.2.1 Packetization of 10 ms

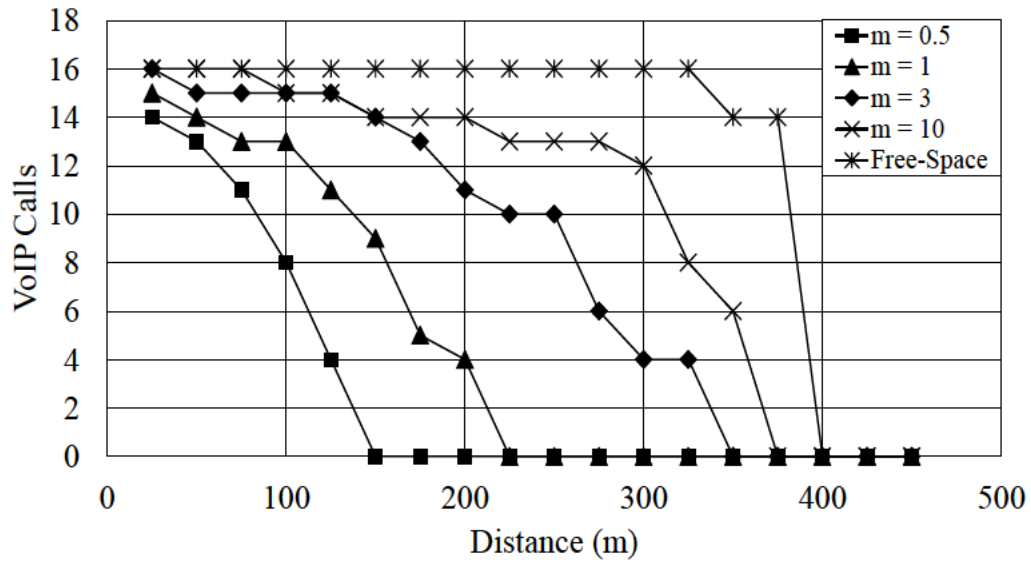
For the G.711 codec, Fig. 3.16 and Fig. 3.17 show how the number of simultaneous valid VoIP calls supported by the network decreases as the distance between VoIP sources and the access point increases for transmission rates of 11 and 54 Mbps, respectively. Different system responses are noticed depending on the propagation model used and the value of the shape factor m , which characterizes the Nakagami- m model as discussed above. During the simulations, m took the following values: 0.5, 1, 3, and 10. As shown in Fig. 3.16 (a), the Free Space model reaches the longest distance (375 m) maintaining the maximum number of valid calls, and suffers a very sudden fall from the maximum to zero in a shorter gap of distance. On the other hand, the Nakagami- m model shows a smoother fall on the capacity, being the work range (i.e., the range accepting valid VoIP calls) much lower than the one obtained for the Free Space model, specifically, 325, 275, 150, and 75 m when m takes values of 10, 3, 1, and 0.5, respectively. Moving wireless nodes to distances greater than those, no valid calls are supported by the system. In order to study the behavior for a transmission rate of 54 Mbps, Fig. 3.17 (a) shows that the distance ranges with valid calls are longer for 54 Mbps than for 11 Mbps for scenarios with severe fading (125 and 200 m for m values of 0.5 and 1, respectively). This behavior is not noticeable when fading level is

less pronounced, i.e., in Free Space and Nakagami- m scenarios with greater values for m . In this case, the same work range is reached using both 11 and 54 Mbps. As expected, the gradual decline on the coverage range observed when decreasing m is related to the Nakagami- m fading channels. Lower values for m represent greater levels of fading, with a negative effect on the working range.

In addition to coverage range, the maximum VoIP capacity of the system is also affected by fading. This effect is studied in the following by analyzing the maximum capacity reached on each scenario in the shortest distance between wireless nodes and



(a) Codec G.711



(b) Codec G.726 (24 Kbps)

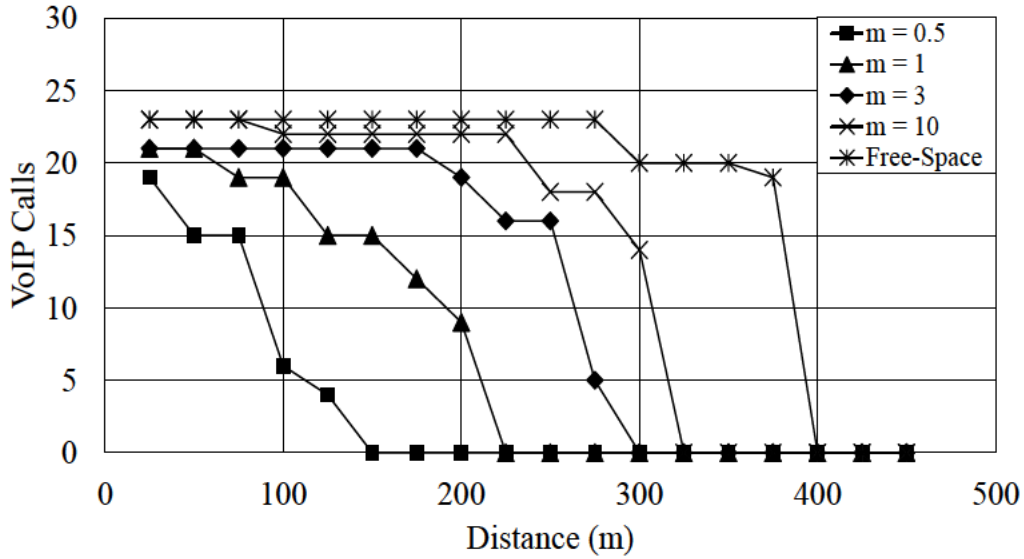
Figure 3.16. VoIP calls supported in the wireless IEEE 802.11g network at 11 Mbps and packetization of 10 ms.

the access point, i.e., 25 m. Scenarios with m taking values of 0.5 or 1 do not reach the same system capacity attained in Free Space scenarios (or in those with greater values for m). All the simulated scenarios, with m equals to 0.5, give support to (at least) one call less than the other scenarios. Furthermore, the system capacity is not constant with distance, i.e., a progressive fall in the number of valid calls supported by the system is noticed when the distance between the access point and the wireless nodes increased.

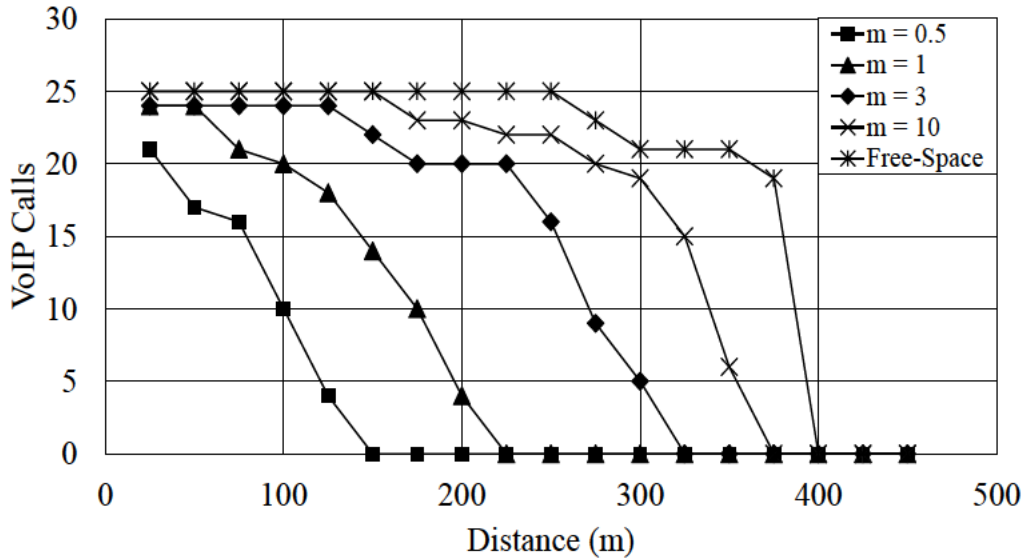
Fig. 3.16 (b) and Fig. 3.17 (b) show the system response when using codec G.726 at a coding rate of 24 Kbps for two different bandwidths of 11 and 54 Mbps, respectively. The same behavior as regards the system working range and capacity is observed. Therefore, no matter the codec used, fading affects the VoIP system in both aspects, (i) the decrease of maximum distance reached with valid calls, and (ii) the maximum number of valid calls accepted by the system.

Regarding the effect of the type of codec on the maximum system capacity, it is noticed a small increase using the low bit-rate codec G.726 compared to the results obtained for G.711. As mentioned above, the maximum number of VoIP calls is assessed taking into account the achieved QoE (MOS), the one-way delay, and the PLR. Accepting a call as valid with a QoE (MOS) limit of 3.1, the following results are obtained. For Free Space model at 54 Mbps, G.711 (64 Kbps) achieves 23 valid calls against 25 calls for G.726 (24 Kbps). Using the most severe fading scenario ($m = 0.5$), only 19 and 21 calls, respectively, are accepted as valid as shown in Fig 2 (b and d). On the other hand, at 11 Mbps, 15 valid calls are achieved for G.711 and 16 calls for G.726 (24 Kbps), for no-fading Free Space scenario. However, in a Nakagami- m environment, with $m = 0.5$, 13 calls for G.711 and 14 calls for G.726 (24 Kbps) are achieved (see Fig. 3.16). It could be expected a greater difference between G.711 and G.726 (24 Kbps) in the number of valid calls accepted. As discussed in previous sections, the small increase in the capacity obtained using a low bit-rate codec is due to the legacy overheads and preambles introduced by 802.11g, which mask the effect of low bit-rate codecs.

Concerning quality in the communications, the effect of the type of codec and the fading channels over the QoE (MOS) are next studied. The maximum QoE (MOS) reached for each call decreased when low bit-rate codecs are used in comparison with the QoE (MOS) obtained for G.711. Focusing on the previously analyzed codecs, i.e., G.711 and G.726 with a coding rate of 24 Kbps, a comparison between the maximum



(a) Codec G.711



(b) Codec G.726 (24 Kbps)

Figure 3.17. VoIP calls supported in the wireless IEEE 802.11g network at 54 Mbps and packetization of 10 ms.

number of VoIP calls accepted by the system and the maximum QoE (MOS) value attained (for the different propagation models and transmission rates used) is shown in Fig. 3.18. Although the capacity reached was greater using codec G.726 (as discussed above), observe how the QoE (MOS) attained in both cases show a big difference; for instance, QoE (MOS) of 4.35 for G.711 and QoE (MOS) of 3.3 for G.726 (24 Kbps), at 11 Mbps, when $m = 0.5$ (see Fig. 3.18 (a)). The quality of the VoIP calls is not only affected by the used codec, but also by the effect of fading channels. As shown in Fig. 3.18 (b), the maximum QoE (MOS) value reached for codec G.726 decreases with lower values of m , i.e., with higher level of fading. On the other hand, codec G.711

shows a greater strength against fading, and almost the maximum QoE (MOS) value of 4.5 is reached in all scenarios.

A comparison between the capacity and QoE (MOS) reached for the three coding rates employed for codec G.726 is shown in Fig. 3.19. The capacity does not exhibit a big increase regarding coding rates. It is worth to mention the results obtained for 54 Mbps (Fig. 3.19 (b)). Observe how the attained capacity using a coding rate of 24 Kbps is lower than the ones obtained with the other coding rates. This is related to the low QoE (MOS) value assessed for each call, which makes us to discard active calls because their level of quality are lower than the QoE (MOS) limit established, i.e., 3.1.

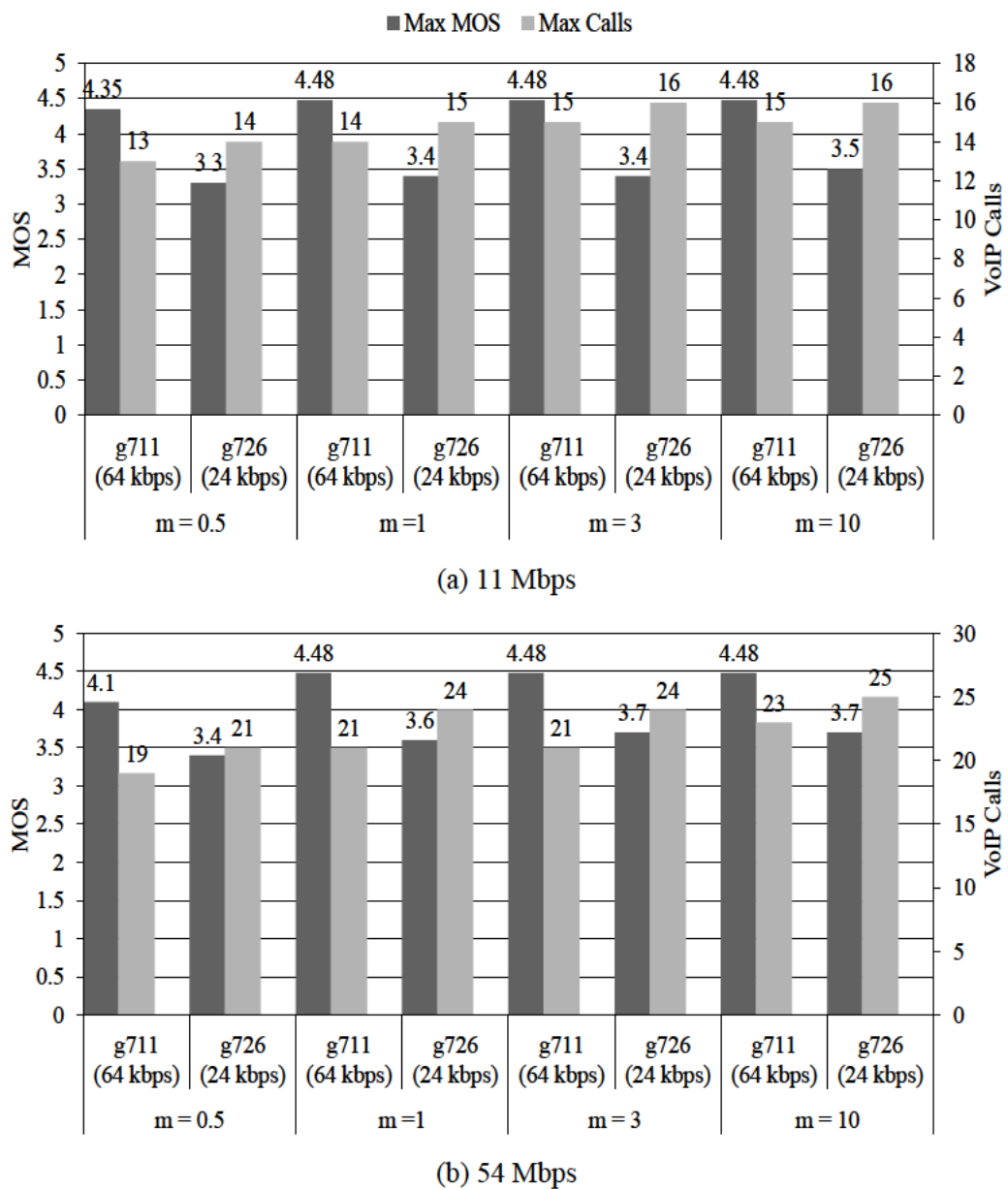
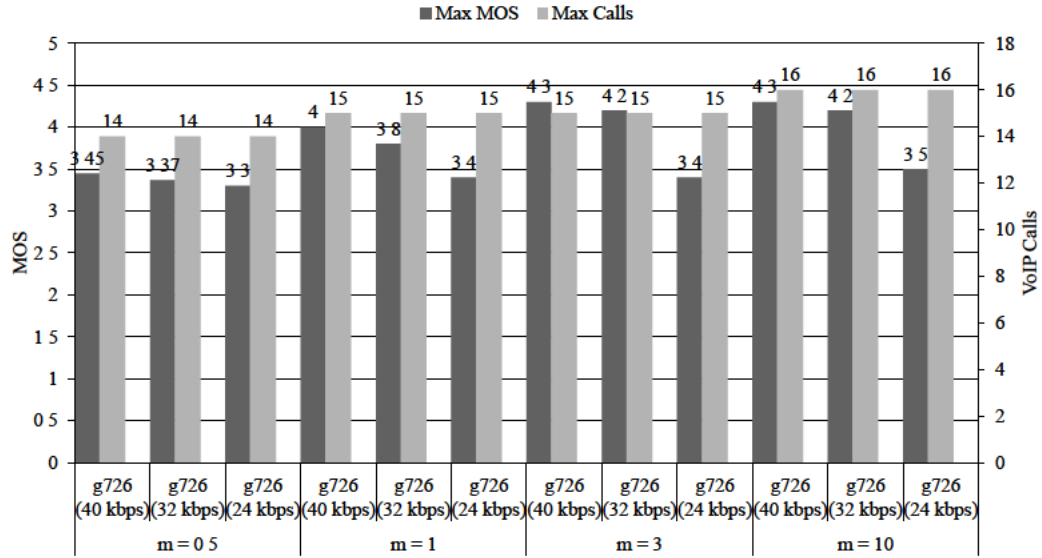
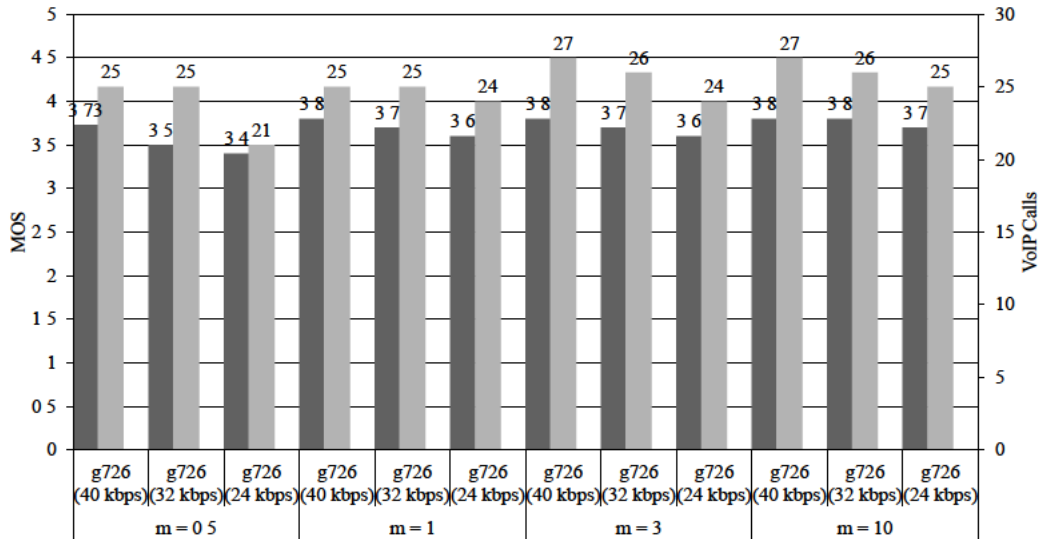


Figure 3.18. Comparison of maximum VoIP capacity and QoE (MOS) obtained in the VoIP calls with packetization of 10 ms and codecs G.711 A-law and G.726 (24 Kbps) for the different propagation models under study.



(a) 11 Mbps



(b) 54 Mbps

Figure 3.19. Comparison of maximum VoIP capacity and QoE (MOS) obtained in the VoIP calls with packetization of 10 ms and codec G.726 at 24, 32, and 40 Kbps for the different scenarios under study.

In order to study the effect of fading channels over the key network metrics affecting the communication quality, the one-way delay and PLR are analyzed in the following. Thus, Fig. 3.20 depicts the evolution of QoE (MOS) and delay as the number of calls accessing the system increases. As in the previous section, these metrics are evaluated in the transition distance for each scenario, i.e., when the system capacity decays from its maximum to lower levels. As mentioned previously, a maximum one-way delay limit to 80 ms was set. Observe that, when delay exceeds that value, the QoE (MOS) drops under 3.1, so VoIP calls are not accepted as valid ones (Fig. 3.20). The same behavior is detected by analyzing the PLR. A PLR limit of 5% was set, meaning that

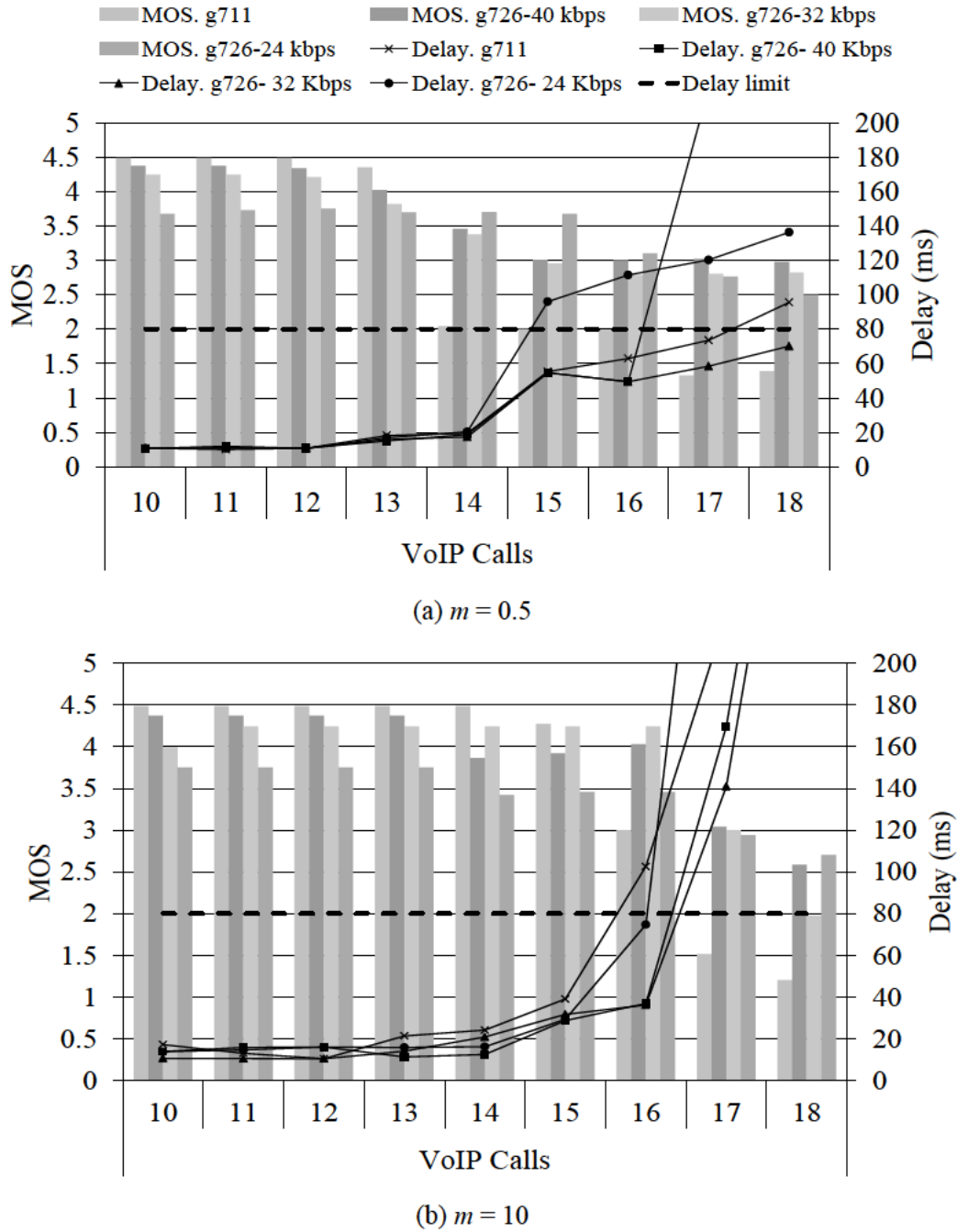


Figure 3.20. QoE (MOS) and delay measured in Nakagami- m scenarios for 11 Mbps and a packetization of 10 ms. For each group of bars, the bar sequence is as follows, G.711, G.726 (40 Kbps), G.726 (32 Kbps), and G.726 (24 Kbps).

communications that suffer rates over this limit should not be accepted as valid. As shown in Fig. 3.21, when the PLR exceeds that threshold, quality estimation falls below 3.1, so this number of calls is set as unacceptable. Therefore, a call can be set as invalid if just one of the parameters, delay or PLR, reaches its threshold. For instance, let us analyze the case of codec G.726 at 40 Kbps. When $m = 10$ at 11 Mbps, it reaches a maximum capacity of 16 calls. Observing Fig. 3.20 (b), notice that the delay limit is

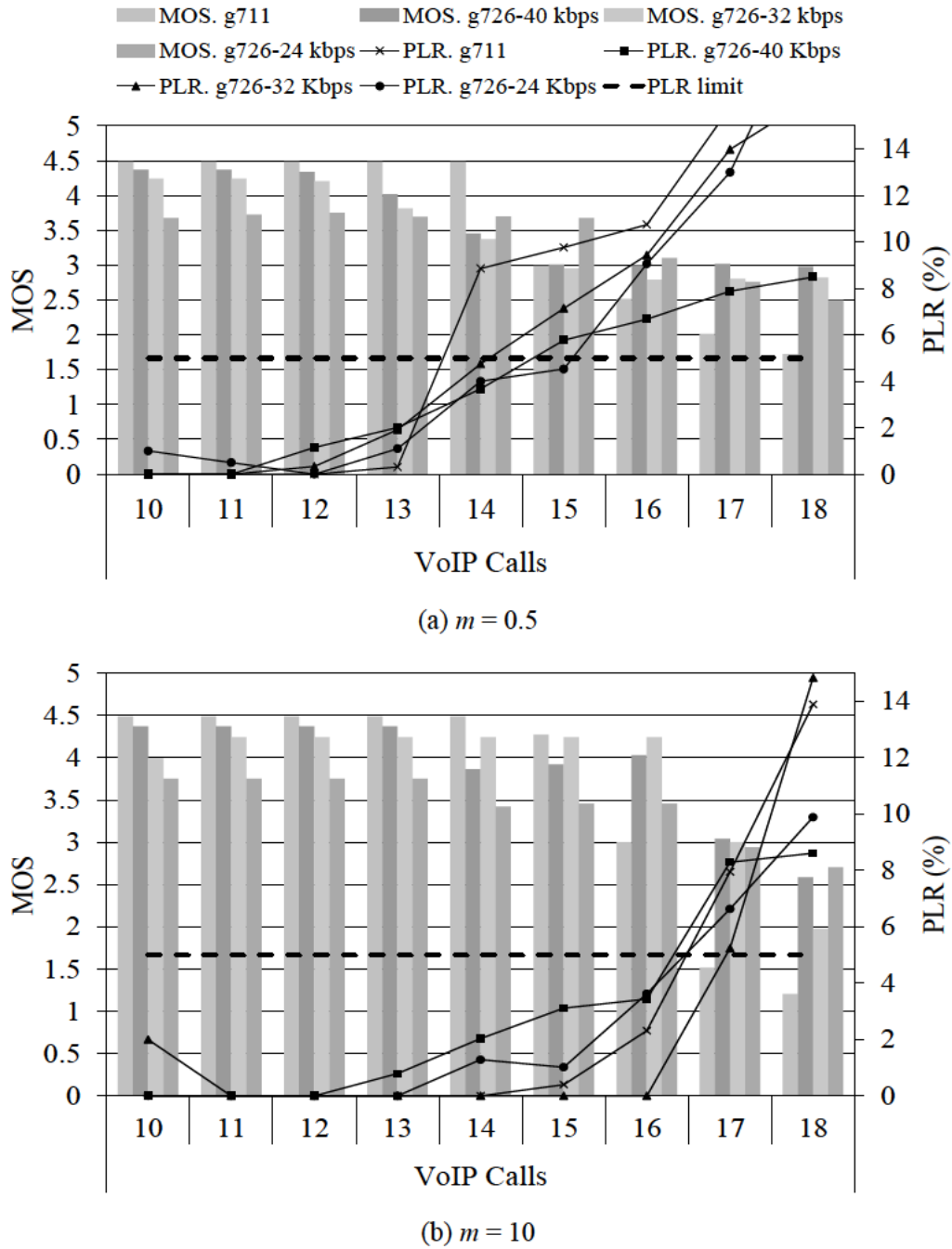


Figure 3.21. QoS (MOS) and PLR measured in Nakagami- m scenarios for 11 Mbps and a packetization of 10 ms. For each group of bars, the bar sequence is as follows, G.711, G.726 (40 Kbps), G.726 (32 Kbps), and G.726 (24 Kbps).

overpassed with 17 calls on the system, rising from 36.94 ms (16 calls) to 169.7 ms (17 calls). A similar behavior holds for PLR (Fig. 3.21 (b)), achieving a PLR of 3.43% for 16 calls and 8.28% for 17 calls. In addition, the QoS (MOS) value for 17 calls is 3.0, so that, this number of calls is taken as not supported by the system.

TABLE 3.6. Capacity of the system for the Free Space model obtained through simulation (in VoIP calls)

Codec	TX Rate			
	11 Mbps		54 Mbps	
	Packetization Interval		Packetization Interval	
	10 ms	20 ms	10 ms	20 ms
G.711	15	27	23	44
G.726 (40 kbps)	16	29	27	51
G.726 (32 kbps)	16	29	27	51
G.726 (24 kbps)	16	29	25	50

3.3.4.2.2 Packetization of 20 ms

In order to study the effect of fading channels on larger packets, a 20 ms packetization was also employed in the simulations. Referring to the drop of the system capacity with the distance, the same behavior as for the lower packetization interval is noticed, i.e., scenarios with lower values of m support less number of valid calls and their coverage range decreased. In addition, due to the increase on the voice payload, an increment on the system capacity is observed in both, fading and non-fading scenarios. Table 3.6 shows the results obtained for the Free Space environment. Observe the difference of capacity attained between codec G.711 and G.726 using a packetization interval of 20 ms at 54 Mbps. In contrast with what happened using a packetization of 10 ms, a remarkable difference is obtained due to the legacy preambles and overheads loss weight in front of the voice payload.

Regarding the quality achieved for each codec under evaluation, it is obtained the same levels of quality as for the packetization of 10 ms, again, showing a decrease in scenarios with greater level of fading. Comparing the overall performance between the two packetization time-lengths, the 20 ms packetization shows more profitable results, because without decreasing the level of quality the capacity almost doubles compared with the 10 ms packetization.

3.4 Conclusion

In this chapter, the performance of IEEE 802.11 infrastructure networks supporting VoIP traffic has been evaluated from a QoE perspective. In order to provide results as realistic as possible, the impact of fading channels on the wireless transmission system has been taken into account. Thus, four different propagation models, namely, Rice, Rayleigh, Nakagami- m , and Free Space, have been employed. By means of these diverse physical layer characterizations it has been simulated several transmission-

channel conditions, demonstrating that fading has a severe impact on (i) the maximum system capacity, (ii) coverage range, and (iii) QoE of each VoIP call. It has been also examined different bandwidths, coding schemes, and packetization time-intervals to evaluate the system behavior with different combination of these parameters under the effect of fading channels. From the attained results, it was obtained that the most hostile scenarios supported noticeably less VoIP calls than non-fading scenarios, reaching lower working ranges with differences ranging from 100-200 m, too. It was also showed that using low bit-rate codecs (G.726) allows an increase on the system the capacity and coverage range, compared with those attained using a no-compression codec (G.711). However, a severe drop on the QoE of the VoIP calls was noticed when employing the low bit-rate codec in comparison with the QoE (MOS) levels obtained using no-compression codec. Finally, the impact of setting different packetization time-intervals was also evaluated, obtaining advantageous results for the packetization of 20 ms because, without decreasing the level of quality, the capacity almost doubles compared with the 10 ms packetization.

To sum up, it can be concluded that the IEEE 802.11 infrastructure networks are capable to support highly demanding traffic such as VoIP communications with good levels of quality. However, the impact of the wireless transmission environment should not be ignored, as it has a great influence on the communications quality.

3.5 Contributions Related to this Chapter

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “On the impact of Rice and Rayleigh channel characterization on VoIP services: a QoE approach,” *Recent Advances in Communications and Networking Technology*, vol. 2, no. 2, pp. 82 – 90, 2013.

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “QoE simulation analysis of VoIP traffic support under Nakagami-m fading channels,” *Signals and Telecommunication Journal*, vol. 2, no. 2, pp. 62 – 73, 2013.

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “On the effect of the physical layer on VoIP Quality of user Experience in wireless networks,” in *IEEE International Conference on Communications (IEEE ICC'13) - 3rd IEEE International Workshop on Smart Communication Protocols and Algorithms (SCPA 2013)*, 2013, pp. 1056 – 1060.

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “Performance evaluation of QoE in VoIP traffic under fading channels,” in *World Congress on Computer and Information Technologies (WCCIT'13) - The International Conference on Mobile Applications and Security Management (ICMASM'2013)*, 2013, pp. 151 – 152.

R. Sanchez-Iborra and M.-D. Cano, “Evaluación del impacto de los canales de desvanecimiento sobre la calidad de experiencia en sistemas VoIP,” *Jornada de Introducción a la Investigación de la UPCT*, vol. 6, no. 1, pp. 103 – 105, 2013.

Chapter 4

Performance Evaluation of 802.11g Mobile Ad-hoc Networks Supporting Multimedia Traffic: a QoE Perspective

4.1 Introduction

Currently, Mobile Ad-hoc NETWORKs, also known as MANETs, are one of the wireless network-access topologies receiving more attention by the research community. As explained in previous chapters, the ad-hoc mode is another network configuration-scheme defined in the IEEE 802.11 standard. This configuration mode does not employ a central access point and all the nodes composing the network assume both host and router roles, helping each other to relay packets in their path to the destination. That feature allows the quick deployment of self-configurable networks without a pre-existent infrastructure. This type of network is considered of great interest due to its multiple benefits such as network coverage extension, scalability, and mobility capabilities provided to the nodes. However, MANETs pose additional challenges, especially related to node motion. Thus, they should be able to auto-configure themselves, reacting to changes in their topology. Moreover, as discussed in previous chapters, the transmission of multimedia traffic entails real-time constraints, necessary to provide these services with acceptable levels of quality. For that reason, efficient routing protocols are needed, capable of setting the paths between nodes as fast as possible and with the ability of reacting to scenario changes.

One routing protocol that has been devoted a notable research effort during the last years is the Better Approach To Mobile Ad-hoc Networking (BATMAN) protocol. BATMAN is an open-source proactive protocol designed to operate in multi-hop ad-hoc mesh networks. Currently, it is defined as an IETF's draft [204], and its main contribution is that nodes in the network do not try to determine the whole path to each destination; instead, BATMAN nodes only know the best next-hop to a given destination without routing-information exchange among nodes. With this feature, BATMAN aims to outperform other widely-used routing protocols such as OLSR (Optimized Link State Routing), AODV (Ad-hoc On-Demand Distance Vector

Routing), or DSDV (Destination-Sequenced Distance Vector) [205], [206], by employing a simpler algorithm, which consumes less CPU resources, and, therefore, allowing battery saving. Green networking is an emerging trend, focused on the development of environmental-friendly architectures by reducing the power consumption of network devices, and it stands also as a key factor for the deployment of long lifetime wireless technologies such as MANETs and Wireless Sensor Networks (WSN).

This chapter presents an evaluation of the performance of the BATMAN ad-hoc routing protocol [204], supporting multimedia traffic (both VoIP and video-streaming). The performance of BATMAN is compared with that obtained by the deeply studied OLSR (Optimized Link State Routing) routing algorithm [207]. This protocol, whose status is denoted as experimental RFC by the IETF, has shown better performance than other proactive routing protocols in multimedia-delivery scenarios [208] and a number of works analyzing its pro and cons can be found in the literature. However, to the best of author's knowledge, it has not been published a performance evaluation of both protocols in a real test-bed composed by a significant number of nodes. Additionally, most of the works addressing quality attained by multimedia services over MANET scenarios obtained their results either only by simulation or only by using small real test-beds, separately. This is an important point, as the simulation studies should be corroborated by experiments in realistic scenarios in order to evaluate the correlation between the results obtained by both methodologies. For those reasons, some of the results of this chapter are obtained through both computer simulation and from a real test-bench by making use of the open-access platform Emulab [9] for a considerable number of nodes composing the network. All results are provided in terms of both QoE and QoS.

In the following, a wide description of the routing protocols evaluated in this chapter, namely, BATMAN and OLSR, is provided. Then, the two different-nature test-benches under consideration are explored. Finally, the obtained results are presented and discussed. For the sake of clarity, the attained outcomes are separated attending to the type of the evaluated service, VoIP or video-streaming.

4.2 Routing Protocols in MANETs

Due to the dynamic nature of MANET scenarios, efficient multi-hop routing protocols are needed. One of the most extended classifications to sort ad-hoc routing protocols is to differentiate them into proactive, reactive, or hybrid protocols.

Proactive protocols, also known as table-driven protocols, maintain the routing information even before it is needed, i.e., every node in the network keeps routing information to every other node. Therefore, proactive protocols are not suitable for large-scale networks, as each node needs to maintain node entries in their routing tables for every element in the network, and much control information is needed, which causes extra bandwidth consumption. On the other hand, the routing information is always up-to-date and the nodes are able to retransmit immediately an incoming packet to the destination next hop, hence reducing network delays which suppose an improving on time-sensitive services and applications. Examples of proactive protocols are BATMAN, OLSR, or DSDV protocols.

The reactive or on-demand routing protocols only calculate the route when it is required. In other words, there is no routing information maintained or even routing activity at the network nodes if there is no communication; if a node has to send a packet to other one, then the route is looked-up on-demand and the connection is established. That type of operation reduces the bandwidth consumption but increases the latency when a packet has to be retransmitted. AODV [209], DYnamic Manet On-demand routing (DYMO) [210], or Temporally-Ordered Routing Algorithm (TORA) [211] are examples of such reactive protocols.

Finally, hybrid protocols combine the advantages of proactive and reactive algorithms. The routing is initially established with some proactively prospected routes and then serves the demand from additionally activated nodes through reactive flooding. An example of this category is the Zone Routing Protocol (ZRP) [212].

Due to the tough requirements imposed by multimedia services, the routing-protocols efficiency is crucial in order to provide these services with high quality. In a MANET context, it is essential reacting to network changes, finding dynamically the best routes by introducing as less delay and packet loss to the multimedia stream as possible. Next, it is presented a brief description of the routing protocols under study: OLSR and BATMAN. The election of these algorithms is aimed at fairly comparing the

performance of BATMAN, which is a promising up-to-date proactive algorithm, with the yield of one of the most extended proactive routing protocols: OLSR.

4.2.1 OLSR

This routing algorithm for MANETs is a well-known IETF's experimental-draft [207], based on the traditional concept of link-state routing algorithm. OLSR makes use of two different types of control packets, namely, HELLO and Topology Control (TC) packets. HELLO packets are employed by nodes to find out their neighboring nodes, so these packets are not retransmitted to the entire network. Once each node knows its "neighborhood", it starts sending TC packets including its neighbors and the state of the links established between them. This helps other nodes to build the network topology. The improvement introduced by OLSR is that the amount of control traffic, specifically TC packets, in the network is reduced by employing the Multi-Point Relaying (MPR) strategy. To do this, each node selects a set of its neighbors to retransmit its TC packets, including them in a field within the packet header; the remaining nodes can read these packets but are not allowed to retransmit them. With a proper strategy selecting the retransmitter neighbors, all the destinations are reachable by all the nodes in the system without flooding the entire network. Additionally, TC packets include a sequence number in order to avoid infinite retransmissions due to undesirable loops.

Using different simulation frameworks, several works have also addressed comparative performance evaluations of other ad-hoc routing protocols against OLSR supporting VoIP and video traffic [213], [214], but not including BATMAN. Adam *et al.* [213] presented a comparison among OLSR, AODV, and DSR (Dynamic Source Routing) [215]. In this work, authors carried out a performance evaluation of these protocols supporting CBR (Constant Bit Rate) traffic simulating a multimedia service. They found out that AODV beats OLSR in terms of jitter and Packet Delivery Ratio (PDR), but at the expense of higher routing overhead. On the other hand, OLSR showed better performance than AODV in terms of delay for a high number of simultaneous connections in the network. In [214], the performance of OLSR supporting CBR traffic was also compared with that obtained by the routing Protocol for Unified Multicasting Through Announcements (PUMA) [216]. Several QoS metrics, namely, throughput, latency, and PLR were assessed. In every performed experiment, OLSR beat PUMA, obtaining promising results, although just one simultaneous multimedia flow in the

network was considered. The good results in terms of QoS in comparison with other routing algorithms attained by OLSR in these works and others [208] have led us to choose it as the reference scheme to compare BATMAN with.

4.2.2 BATMAN

This proactive protocol is under development by the “Freifunk Community” [217]. The novelty of BATMAN resides in the decentralization of the knowledge about the routes; in other words, single nodes do not have routing tables for the entire network. Instead, each node determines one single-hop neighbor for each destination in the mesh, which can be utilized as the best gateway to communicate with the destination node. Thus, a very fast and efficient routing scheme is developed, creating a network of collective intelligence, and allowing low CPU and consequently less battery consumption for each node [218]. The protocol operation is as follows. On a regular basis every node broadcasts an OriGinator Message (OGM), thereby informing its link-local neighbors about its existence. Link-local neighbors which are receiving the OGM messages are relaying them by rebroadcasting, according to the specific BATMAN forwarding rules. The BATMAN mesh network is therefore flooded with OGM messages until every node has received each of them at least once, or until they got lost due to packet loss occurred in the communication links, or until their TTL value has expired. The number of OGM messages received from a given node via each link-local neighbor is used as base to estimate the quality of a route. In order to be able to find the best route to a particular end node, BATMAN counts the OGM messages received from each node in the network and logs which link-local neighbor relayed the message. Using this information BATMAN maintains a table with the best link-local route towards every other node in the network. By using a sequence number, included in each OGM, BATMAN can distinguish between new OGM packets and their duplicates, ensuring that every OGM gets only counted once. Notice that OGMs should not be taken actually as routing-information interchange packets, as they only act as a “hello packet”, not containing any information about routing tables, link states, etc.

The performance of BATMAN supporting multimedia traffic in real scenarios has been barely evaluated in the related literature. Two relevant works addressing this topic are those presented by Kulla *et al.* [219] and Seither *et al.* [220]. In the former, the performance of BATMAN supporting multimedia traffic (audio and video) in small

test-beds was evaluated. In their results, the authors demonstrated the capability of BATMAN-managed MANETs formed by 5 nodes to support VoIP traffic; however, they also showed some issues to support video-traffic, which is much more bandwidth demanding than VoIP. In turn, work in [220] presented only a QoS comparison between BATMAN and AODV operating in more populated real test-beds. The results showed a noticeably better performance of BATMAN in terms of PLR, throughput, and stability than AODV. However, the traffic patterns employed seem rather limited as the authors made use of ping and constant TCP traffic to test the network.

4.3 Test-Benches

Two different frameworks have been employed to obtain the results showed in the following sections. Similar to the previous chapter, the Omnet++ v4.4.2 simulator was used with the InetManet 2.2 framework [10]. These libraries include the necessary ad-hoc node model, as well as the implementation of the different-layer elements used in these experiments, namely, 802.11, BATMAN, OLSR, traffic generators, etc. On the other hand, it is assumed that it is also necessary to validate the results obtained through simulation by comparing them with measurements taken from real test-benches. Thus, the Emulab platform [9] was selected to run some of the mesh network topologies under study. Emulab is a network test-bench that provides both wired and wireless nodes with high level of Software customization. This platform has shown accurate results and fidelity [221], especially when no delay-emulation, i.e., delays introduced by the management-system to simulate congestion situations or link-introduced delays, is set. This is the case of the system under study, in which additional delays to those induced by the network were not included.

The evaluated scenario was an IEEE 802.11g wireless network at 54 Mbps in ad-hoc mode. Channel access was according to the Distributed Coordination Function (DCF) scheme without employing the RTS/CTS mechanism. Two ad-hoc network topologies were taken under consideration: (i) a static chain of nodes with a variable number of hops between transmitters (TX) and receivers (RX) as shown in Fig. 4.1 (a), and (ii) a mesh topology illustrated in Fig. 4.1 (b), for different scenario dimensions and node densities. Additionally, it was employed three mobility configurations, namely, static, partial, and full dynamic movement. In both topologies, (i) and (ii), the network was loaded with different number of simultaneous multimedia streams in the system. 10

simulation instances with different seeds have been run for every evaluated scenario. Hence, in order to produce our results, the average value for every measured parameter was taken, avoiding non-representative singularities.

As part of the results is obtained from simulated scenarios, it is important modeling the transmission channel as realistic as possible (please, see chapter 3). For that reason, the effect of the fading channels on the wireless transmission medium was introduced in the simulation by characterizing it with the Nakagami- m propagation model. As showed in the previous chapter, this model permits adjusting the level of fading in the wireless transmission medium by tuning its m factor. This characterization for the physical layer has been widely used in the related literature [222], [223] proving its accuracy. After shaping the Nakagami- m model with different m values, the most precise results according to the real ones were obtained by fixing the m value to 5. This m value is consistent with the indoor channel characterization performed by Oestges *et al.* in a similar campus environment for the 2.4 GHz band [224]. In order to study the greatest

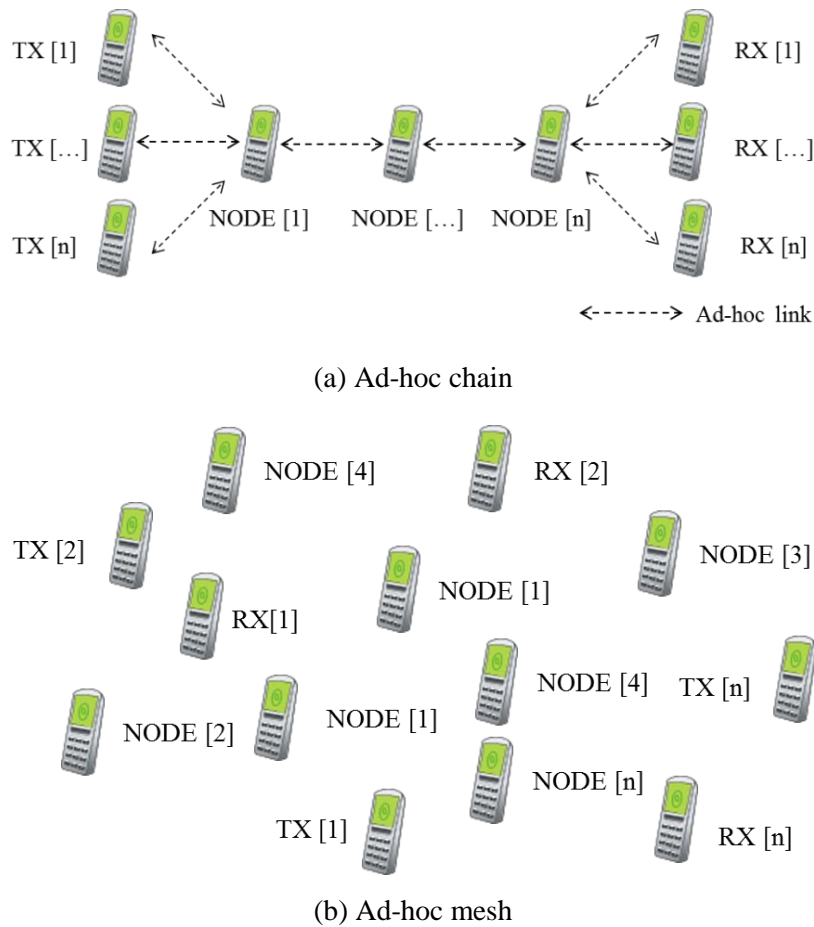


Figure 4.1. Simulation scenarios. TX[n] are the transmitters; RX[n] are the receivers; NODE[n] are the intermediate nodes.

number of hops between source and destination nodes in the evaluated scenarios, the wireless cards transmission power was fixed to 1 mW, since it is the lowest standard value for different-manufacturer wireless cards, such as the Cisco Aironet Series. The wireless cards sensitivity was set to -72 dBm, with a carrier frequency given by the 802.11g standard, i.e., 2.4 GHz, and a SNIR (Signal to Interference and Noise Ratio) threshold of 4 dB.

The test traffic used in the Emulab test-bench was generated by the Distributed-Internet Traffic Generator v2.8.1 [225], which allows sending VoIP traffic following the typical pattern of G.711 A-law codec and video traffic by generating bursty traffic characterized by a given distribution and bit-rate. In the simulation experiment, and as part of the InetManet framework, it was used: (i) the VoIP stream generator, which generates a flow of VoIP packets from an arbitrary sound file using different audio codecs, and (ii) the UDPVideoStreamCli2 and the UDPVideoStreamSvr2 modules. These UDP apps emulate a video-streaming communication in which the video server sends to the client UDP-traffic following a typical pattern for a video transmission, allowing setting the transmission bit-rate. The video packets were fixed to 512 bytes that is a widely employed value for encapsulating video-traffic [226], [227]. Regarding VoIP traffic, it was assumed the same codec as in the real test-bench (G.711) with a VoIP header, i.e., RTP, of 12 bytes and a packet time-length of 20 ms. The starting time for each multimedia communication was generated randomly according to a Poisson distribution in a time interval of (0, 10 s), after the convergence time for each routing protocol. Each multimedia transmission lasted 60 s.

Regarding the quality metrics of interest, both QoS and QoE figures of merit were examined. For video quality evaluation, the latter has been calculated in terms of MOS, from the parametric model presented in [145] and the corrections to this model introduced in [146]. As shown in (4.1), this model allows obtaining accurate video-quality estimations (in terms of MOS) by means of the Packet Loss Rate (PLR), the coding scheme (k), the resolution (a) and the video-coding bit-rate (Br) parameters.

$$V_q = 1 + 4k \left(1 - \frac{1}{1 + \left(\frac{a \cdot Br}{v_1} \right)^{v_2}} \right) e^{-\frac{PLR}{v_3}} \quad (4.1)$$

In (4.1), v_1 , v_2 , and v_3 are experimental factors that depend on the coding scheme, the resolution applied, and the motion characteristics of the video. In this case, as described

TABLE 4.1. 802.11g and routing protocols parameters

Parameters	Bytes	Time
SIFS, DIFS, SLOT (μ s)	-	{10, 28, 9}
CW _{MIN} (slots)	-	31
PLCP preamble (μ s)	-	4
{PLCP, MAC, SNAP} headers (μ s)	-, 28, 8	{16, 4.15, 1.18}
IP + UDP + RTP headers (μ s)	40	5.92
Voice (G.711, 20 ms) (μ s)	160	23.70
Video (μ s)	512	75.85
ACK (μ s)	14	2.07
BATMAN OGM interval (s)		1
BATMAN purge timeout (s)		200
OLSR {HELLO, TC, MID} intervals (s)		{2, 5, 5}
OLSR {Neighbor, Topology} hold times (s)		{6, 15}

above, the widely extended H.264 codec was used for the video-streaming transmission in QCIF format. It was assumed a video-content with low movement, as it is usual in video-telephony calls. Thus, the values assigned to these factors were: $k = 1.12$, $a = 10.8$, $v_1 = 0.366$, $v_2 = 1.32$, and $v_3 = 3.5$ respectively, hence consistent with the figures recommended by the authors [145], [146]. The VoIP QoE (MOS) was obtained using both the E-model and the PESQ model as defined in the ITU-T Rec. G.107 [33] and P.862 [27], respectively. In order to evaluate the protocols performance from a QoS point of view, delay, PLR, and Packet Delivery Ratio (PDR) measurements were also collected. The assigned values for the remaining 802.11g parameters, such as DIFS (DCF Interframe Space), SIFS (Short Interframe Space), slot time, contention window (CW_{MIN}), and different-layers headers are written in Table 4.1, which additionally includes the main routing protocols time-out.

As in the previous chapter, the results section is sub-divided into two parts. First it is evaluated by simulation the effect of the number and density of-ad hoc nodes (nodes per unit of area) on BATMAN performance for a VoIP service. Then, the impact of moving nodes on the VoIP QoE is inspected with the aim of analyzing how BATMAN reacts to topology changes and the influence of the “marginal mobility model” [228]. All the results obtained for BATMAN are compared with those attained by using OLSR. In the second part, by means of the two different-nature test-beds detailed above, it is investigated the capability of BATMAN to support a video-streaming service. Using

both the simulated and the realistic test-benches is an important fact, as the simulation studies should be corroborated by experiments in realistic scenarios in order to evaluate the correlation between the results obtained by both methodologies. Concretely, in this part, the effect of using different bit-rates in video-streaming applications on the QoS/QoE provided when using BATMAN or OLSR in MANETs is assessed. Both routing protocols are analyzed either with their default configuration or with some of their parameters properly adjusted to the network conditions. Finally, a statistical comparison of the results attained in both test-benches is provided, permitting to study the correlation between the results obtained in the simulation framework and those extracted from the realistic test-bed.

4.4 Results

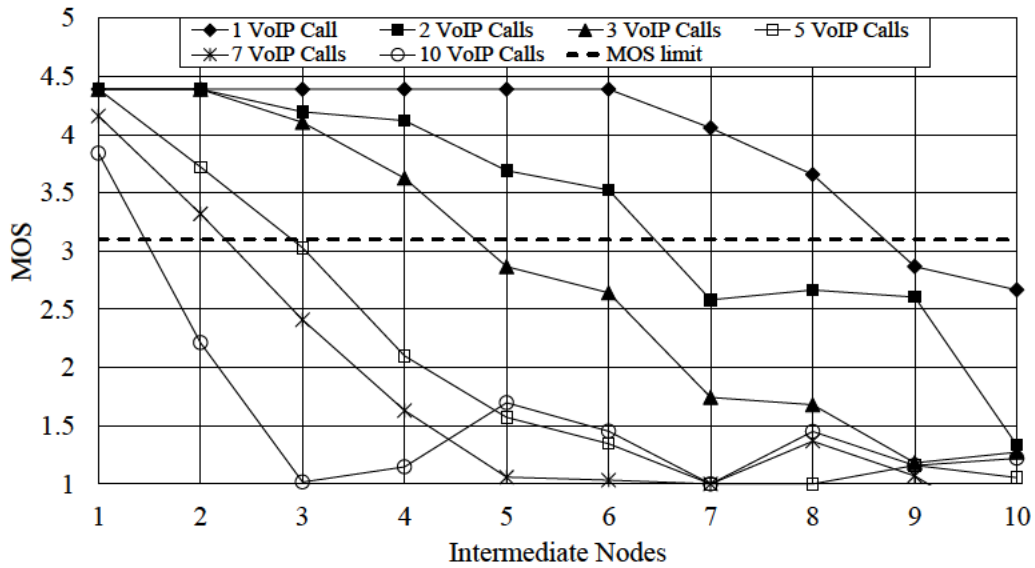
4.4.1 VoIP Traffic

As described above, due to the great sizes of the evaluated networks and the mobility features given to the ad-hoc nodes, the results showed in this section are exclusively extracted from the simulation test-bed. Two different scenarios are taken into consideration: 50 m x 50 m and 125 m x 125 m, with two different numbers of nodes composing the network, namely, 25 and 50 nodes. Regarding nodes movement, it imitates the motion of a pedestrian; each node initiates each simulation at a random position and begins its motion toward a random direction following a linear trajectory with a speed characterized by a Gaussian distribution with a mean value of 1.34 m/s and a standard deviation of 0.26 m/s [229]. As demonstrated in the previous chapter, the impact of the physical layer cannot be ignored when evaluating wireless networks. Thus, the effect of fading channels on the quality of the VoIP calls has been taken into consideration, by comparing the results attained by characterizing the wireless medium with the Nakagami- m model with those obtained in a Free Space situation.

In order to set each individual call as valid, from a QoS/QoE perspective, several metrics are usually used in the bibliography, e.g., QoE (MOS) value, one-way delay, or packet loss probability. Following the guidelines of the ITU-T Rec. G.114 and G.1010, and similarly to the previous chapter, a call is defined as valid if the final QoE (MOS) value attained for this call is over 3.1. This metric allows joining the effect of all the different impairments suffered by the VoIP stream into one single parameter. As discussed above, the QoE (MOS) estimation has been calculated using the PESQ model.

4.4.1.1 Impact of Fading

Fig. 4.2 illustrates a comparison between the average QoE (MOS) achieved for different number of simultaneous VoIP calls under the effect of the physical layer. In this case, it has been used the ad-hoc chain topology shown in Fig. 4.1 (a). The aim of analyzing this topology is to study the number of simultaneous calls that a single node can support and to evaluate the significance of the number of hops in the performance of the VoIP service. The effect of fading channels on the VoIP traffic using OLSR is



(a) Free Space

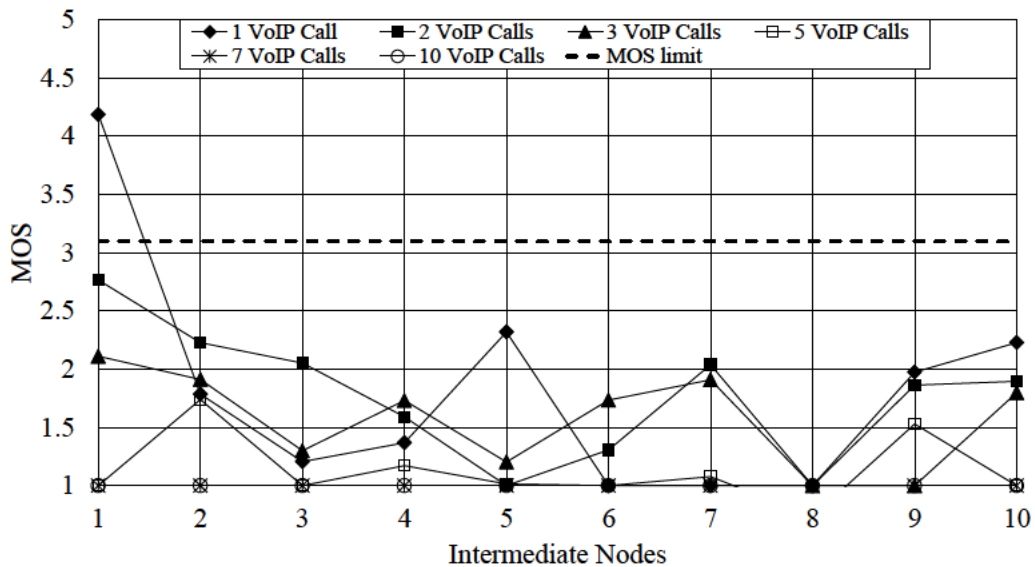
(b) Nakagami- m

Figure 4.2. QoE (MOS) attained for a variable number of VoIP calls and hops between transmitters and receivers. Results for OLSR with Free Space and Nakagami- m propagation models.

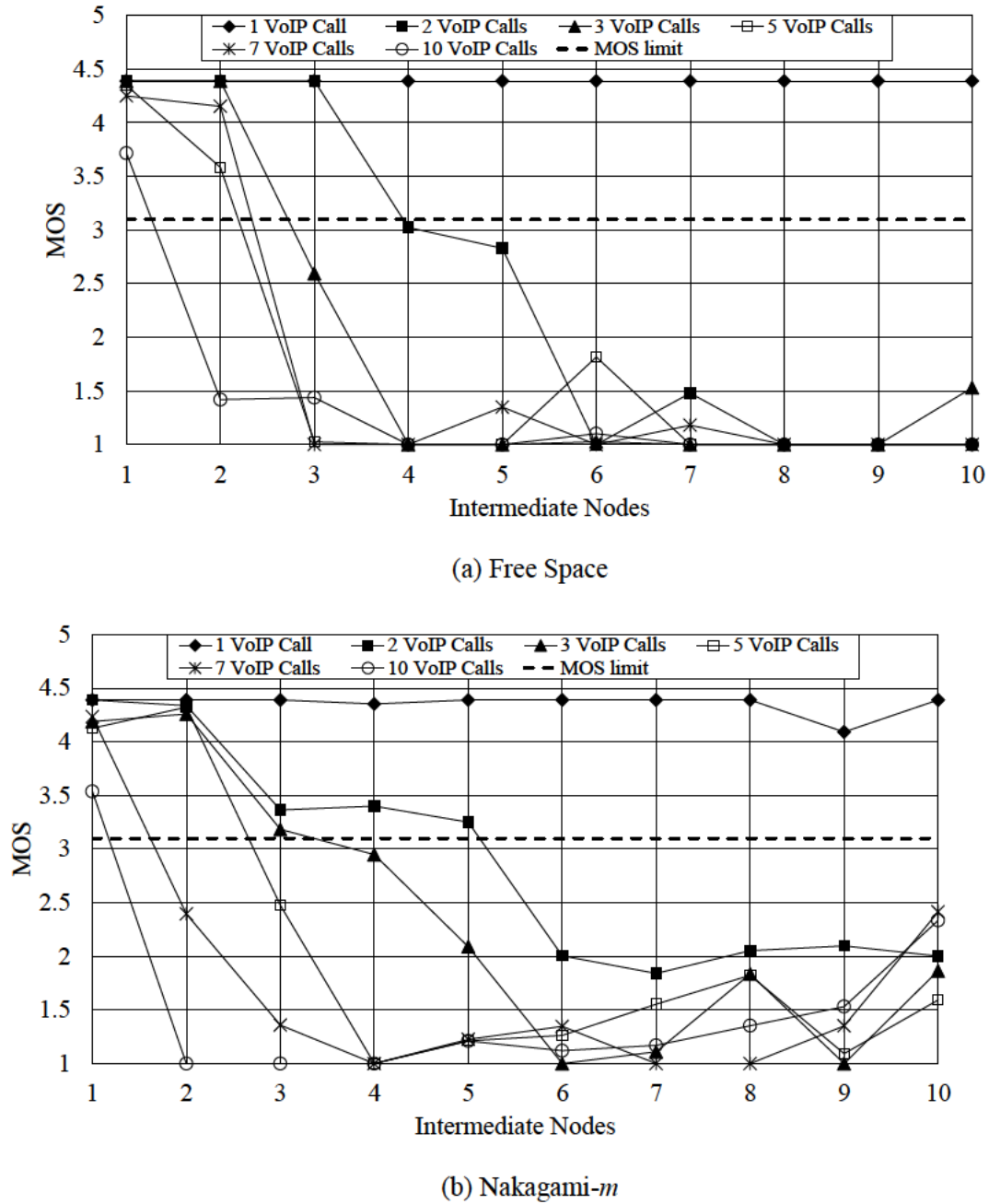
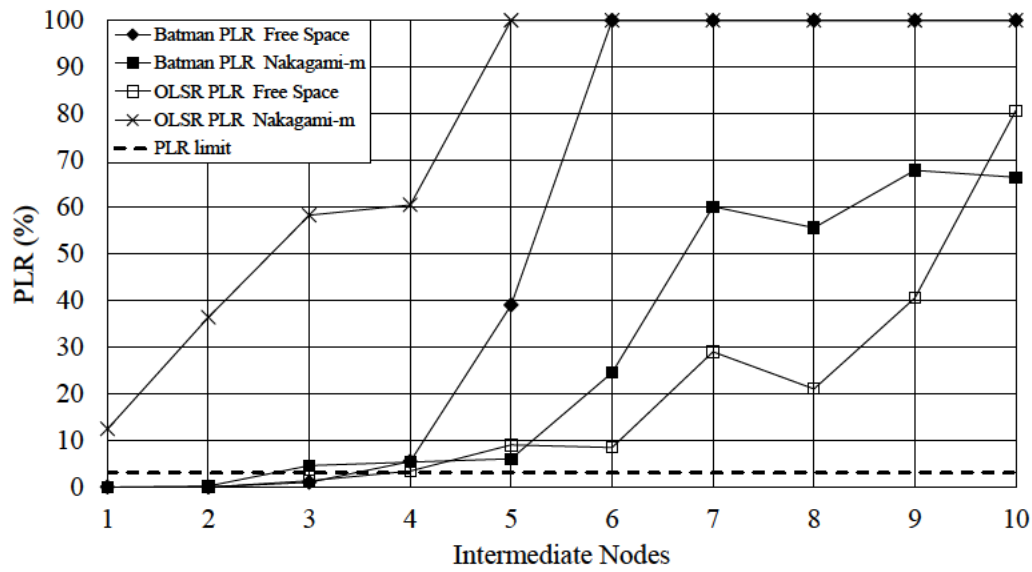


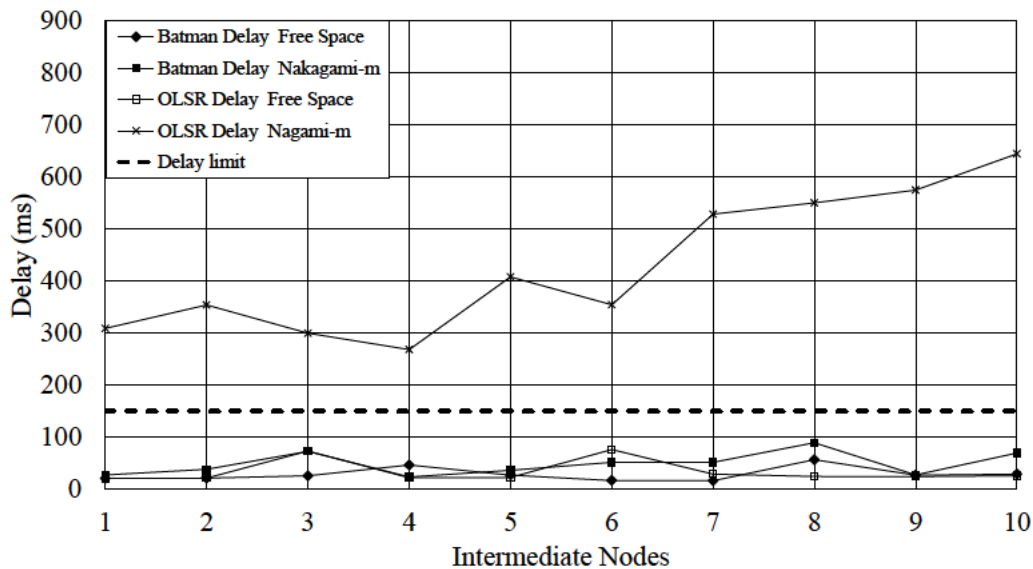
Figure 4.3. QoE (MOS) attained for a variable number of VoIP calls and hops between transmitters and receivers. Results for BATMAN with Free Space and Nakagami- m propagation models.

shown in Fig. 4.2 (a) and 4.2 (b). Observe the severe drop in the QoE (MOS) attained in Nakagami- m scenarios (Fig. 4.2 (b)) compared with the QoE (MOS) values obtained in Free Space environments (Fig. 4.2 (a)): just one call is accepted with one hop between TX and RX, (the rest of the QoE (MOS) values are under 3.1). These results agree with those obtained by Nascimento *et al.* [230] and Singh [231]. In both cases, by employing different propagation models, authors found the same OLSR poor performance in hostile scenarios.

In comparison, the QoE (MOS) fall in fading scenarios is less pronounced using BATMAN (please compare Fig. 4.2 (b) and Fig. 4.3 (b)). BATMAN accepts up to 3 calls with 3 hops between VoIP TXs and RXs. Moreover, 1 and 2 simultaneous calls are established with good levels of QoE (MOS) when up to 10 and up to 5 nodes separate TXs and RXs, respectively. This dramatic degradation in the OLSR network performance is ascribed to the size of its control messages. OLSR makes use of larger packets than BATMAN, so their transmission is more prone to the effect of fading [232], causing an increase in the control packets loss rate, which encumbers the right



(a) Packet Loss Ratio (PLR)

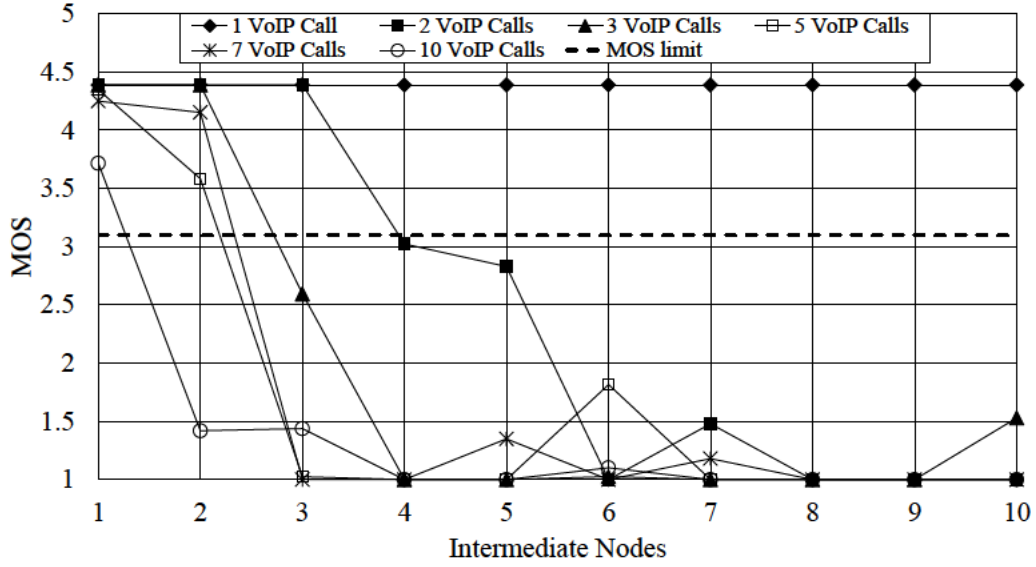


(b) Delay

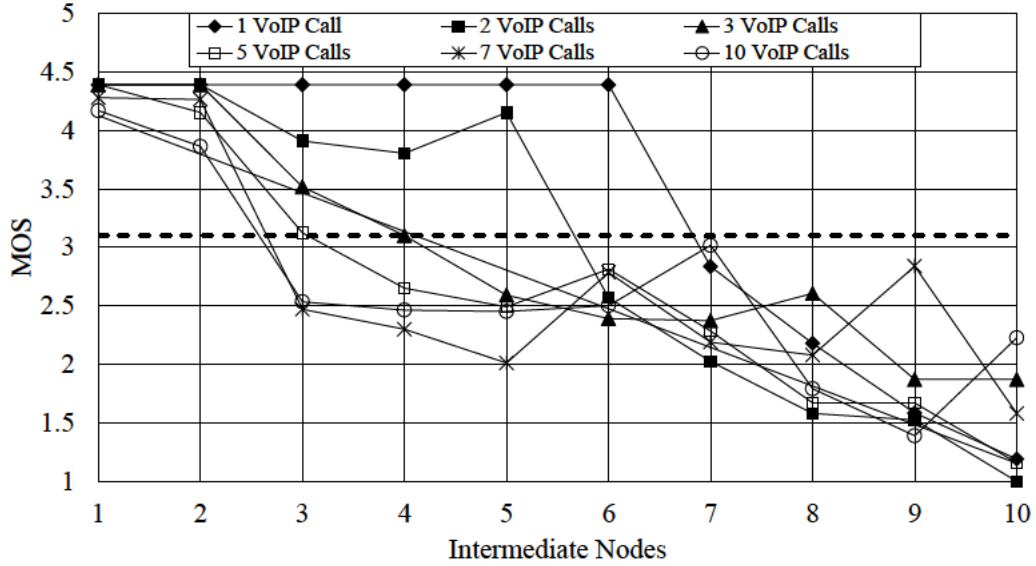
Figure 4.4. PLR (a) and delay (b) suffered by the VoIP streams in chain scenario with two simultaneous calls and variable number of hops between transmitters and receivers, using BATMAN and OLSR.

establishment of the OLSR routes. Hence, the VoIP communications are interrupted or directly cannot be established. In the light of these results, it can be concluded that BATMAN seems to be more robust than OLSR under adverse channel conditions. This is further confirmed from a QoS point of view in Fig. 4.4, which depicts the evolution of the PLR and the delay attained in the VoIP streams for a scenario with two simultaneous calls. Notice the vast rise of packet loss (Fig. 4.4 (a)) and delay (Fig. 4.4 (b)) in fading scenarios managed by OLSR in comparison with those accomplished under Free Space conditions, which explains the deep drop on the QoE (MOS) showed in Fig. 4.2 (b). Also it can be observed in Fig. 4.4 how BATMAN shows its strength to the impact of fading channels, obtaining a more controlled PLR and delay in Nakagami- m scenarios than OLSR.

On the contrary, OLSR over performs BATMAN in the Free Space scenarios (see Fig. 4.2 (a) and 4.3 (a)). Using OLSR, the system accepts up to 2 simultaneous calls in a 6 hops connection and 3 simultaneous calls are supported in transmissions up to 4 hops. However, only one VoIP call is accepted as valid with BATMAN when there are more than 3 hops between TX and RX. Similarly, the system just allows up to 2 simultaneous VoIP calls when the distance between VoIP nodes is 3 hops. Observing Fig. 4.3 (a), Fig. 4.3 (b) and Fig. 4.4, it is clear that BATMAN behaves better in a scenario with fading than in Free Space. This decline in BATMAN performance in terms of QoE (MOS) at Free Space scenarios, as well as in the system capacity in terms of valid number of VoIP calls, is due to the excess of Originator Messages (OGM) generated by the BATMAN algorithm in Free Space, since under these conditions very low packet losses occur. Fig. 4.5 shows a comparison of the QoE (MOS) obtained for the VoIP calls using the default time interval between OGMs (1 s) and that achieved by increasing the OGMs interval up to 2 s in scenarios characterized by the Free Space model. Observe that setting an interval of 2 s, both the number of VoIP calls accepted by the system and the number of hops allowed with a QoE (MOS) over the limit established increase. It is obtained that up to 3 calls are accepted as valid with 4 intermediate nodes, achieving also greater QoE (MOS) values for these accepted calls. That is, the default time interval between OGMs (1 s) produces an increase in both, the transmission-channel occupation and the nodes buffers, with the corresponding growth in packet losses. This behavior is also shown in Fig. 4.6 that describes the evolution of the PLR in a scenario with two VoIP calls (two TX and two RX) and different time intervals between OGMs.



(a) Interval between OGMs: 1 s



(a) Interval between OGMs: 2 s

Figure 4.5. QoE (MOS) obtained for a variable number of VoIP calls and hops between transmitters and receivers using BATMAN with different time intervals between OGMs with Free Space propagation model.

Observe the big difference in the PLR obtained for the different scenarios. As discussed above, using an OGM interval of 1 s, the surplus of OGM messages in the network collapses both the wireless transmission channel and the nodes buffers, increasing the number of VoIP packets dropped. As stated in BATMAN IETF's draft [204], and confirmed by Murray *et al.* [233], BATMAN was not designed to operate on stable and reliable media, so its performance depends on packet loss. This is because of the mentioned broadcasting of OGMs by the nodes, which without packet loss, leads to

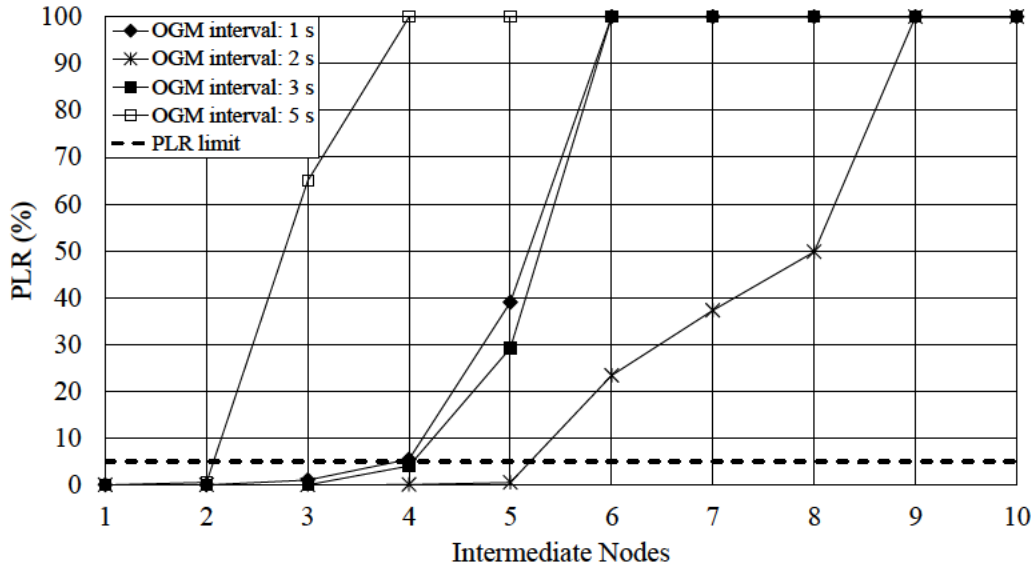


Figure 4.6. Probability of VoIP packet loss in a Free Space chain-network scenario, employing 2 different BATMAN OGM time intervals.

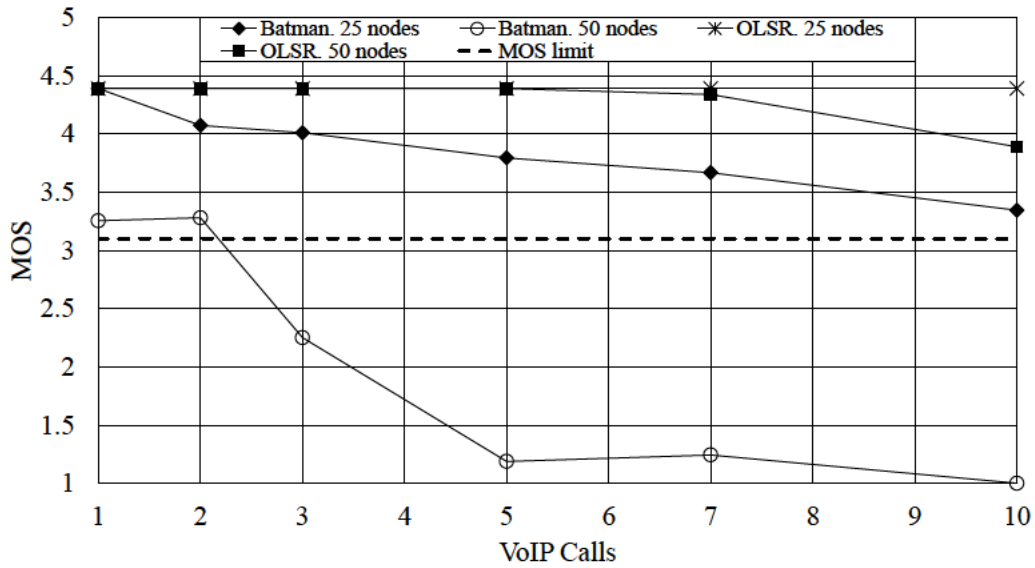
the network saturation. As the Free Space environment does not introduce many packet losses, the OGMs harm the VoIP traffic, resulting in a reduction of the system capacity and VoIP quality. On the other hand, the lack of control messages happened in scenarios with longer OGM intervals, such as 5 s, produces a decrease in the network performance, revealed as an important increase of PLR in the VoIP stream. This behavior is attributed to the speed of convergence of the routing algorithm, which turns very slow for a low number of OGM messages received. It is clear that, for this scenario, an OGM interval of 2 s provides the lowest PLR on the VoIP flows. Therefore, it can be concluded that it is necessary tuning the OGM interval, adapting it to the current network characteristics and considering the trade-off between the number of OGM packets in the network and the lack of enough information in the nodes to establish the routes. This topic will be addressed in section 5.6.1.2.

4.4.1.2 Impact of Number and Density of Nodes

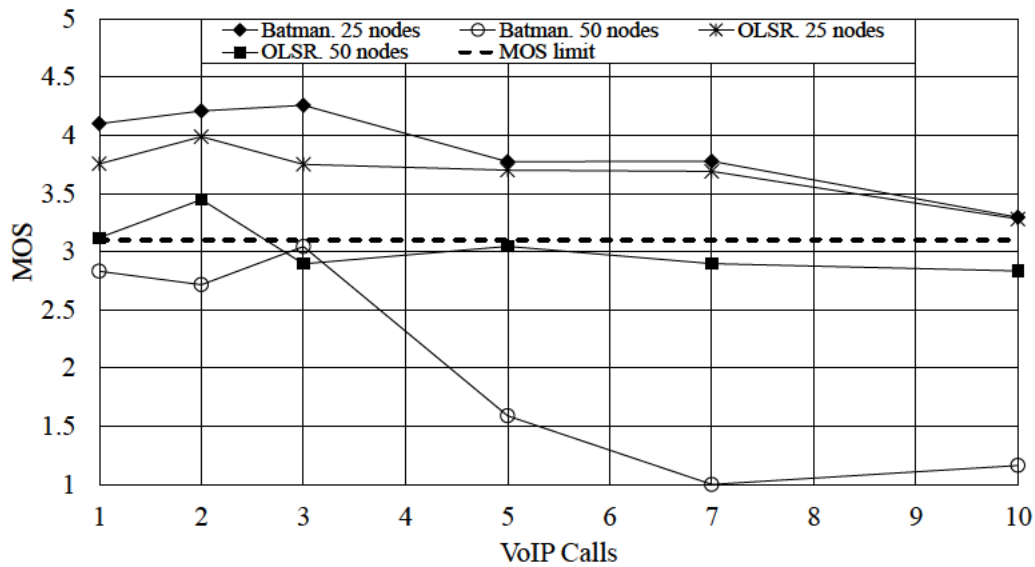
Focusing on the effect of the number of nodes composing the network on the QoE performance, it has been carried out a set of simulations configuring the topology as a mesh network and locating the nodes at fixed random positions. Two different square areas of 50 m x 50 m and 125 m x 125 m are used. In addition, it is also considered two different numbers of nodes, namely, 25 and 50 nodes.

Fig. 4.7 shows the evolution of the average QoE (MOS) in the 50 m x 50 m scenario for two different densities of nodes, two different propagation models, and the two

routing protocols under study. Using BATMAN, it is noticeable an important drop on the quality for the 50-node topology. Observe that, regardless the number of simultaneous VoIP calls in the system, the configuration with fewer nodes always achieves better results. In addition, better results are always obtained using OLSR than using BATMAN in topologies with a larger number of nodes; even in fading environments, which have a deep impact on the performance of OLSR as discussed above.



(a) Free Space



(b) Nakagami-m

Figure 4.7. QoE (MOS) evolution for different node population in the 50 m x 50 m scenario in Free Space and Nakagami- m environments.

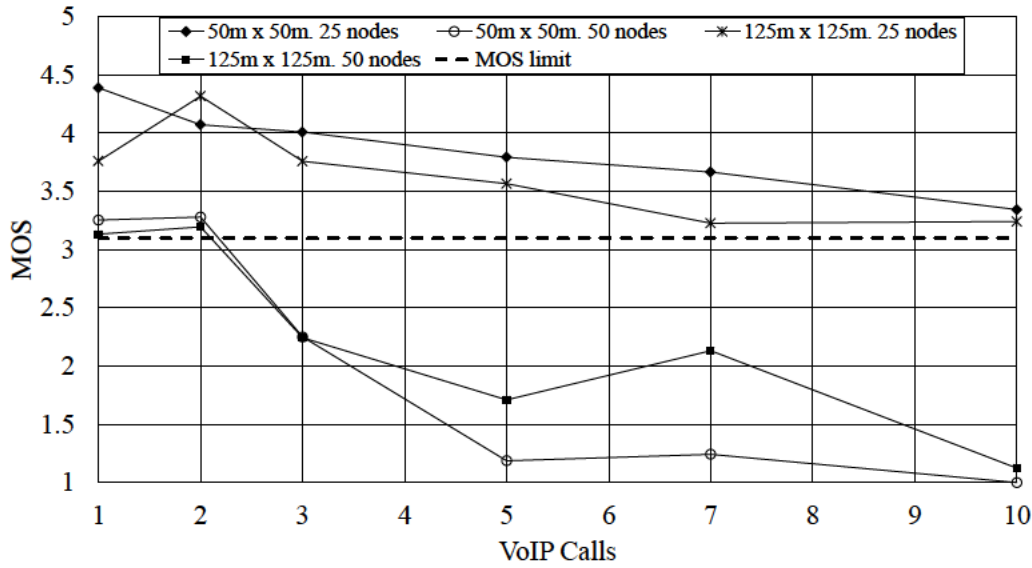


Figure 4.8. Comparison between the QoE (MOS) reached using BATMAN in different-sized Free Space scenarios.

On the other hand, in the 25-node scenario, BATMAN reaches a slightly lower level of quality than OLSR in Free Space situations (Fig. 4.7 (a)), and obtains significantly better results in fading scenarios (Fig. 4.7 (b)). Fig. 4.8 represents the average QoE (MOS) attained with BATMAN for several system loads and different-sized Free Space scenarios with a variable number of nodes composing the network. Observe that, as mentioned above, the QoE (MOS) reached in topologies with 25 nodes is greater than that obtained when the network consists of 50 nodes, regardless the size of the scenario. In the 25-node topologies, 10 simultaneous VoIP calls are supported by the network, with QoE (MOS) levels around 3.3. On the other hand, in more crowded scenarios, just 2 calls are supported by the network satisfying the minimum established level of QoE (MOS). Furthermore, focusing on the network composed by 25 nodes, there is a little decrease in the estimated QoE (MOS) for the largest scenario (125 m x 125 m) comparing with the one obtained in the 50 m x 50 m case. This fact could be explained because of the BATMAN path-finding algorithm. Fig. 4.9 shows a comparison between the average number of hops in the path between the VoIP transmitters and the VoIP receivers in a static mesh network with random node positions. There is a notable difference between the number of hops in the routes calculated by BATMAN and those computed by OLSR, particularly under high density conditions, e.g., in 50 m x 50 m scenarios. The routes obtained by OLSR are always shorter than those established by BATMAN. Additionally, it is found in the simulations conducted for the chain scenario (Fig. 4.1 (a)) that, even with this simple configuration, the number of hops calculated by

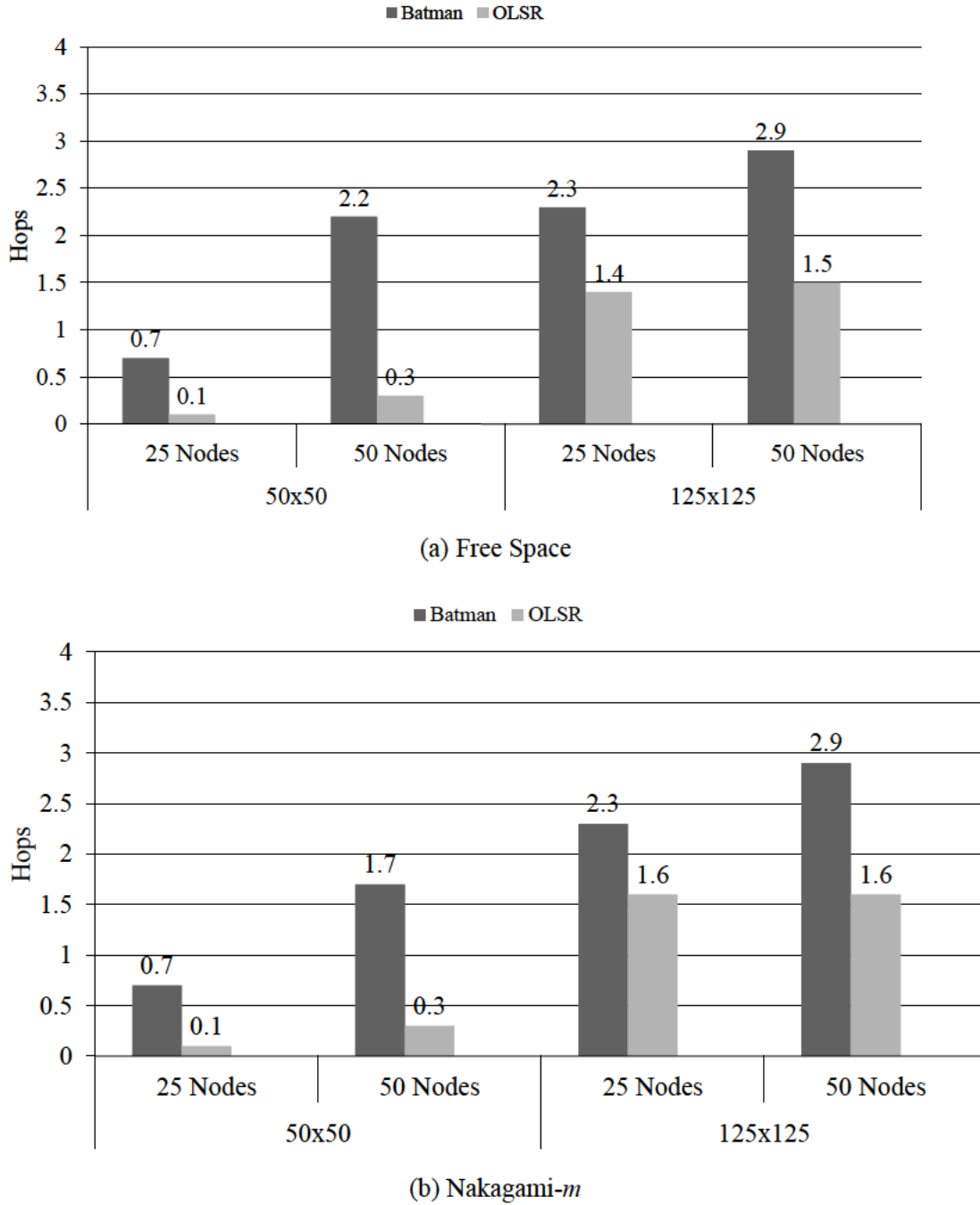


Figure 4.9. Average number of hops between VoIP TXs and RXs in Free Space and Nakagami-*m* environments.

OLSR is lower than that computed by BATMAN. Particularly, OLSR always obtains the best possible result, i.e., the number of hops between TXs and RXs was exactly equal to the number of intermediate nodes in the chain network. However, using BATMAN with several VoIP TXs in the system, some of them use each other as an intermediate node; thus, the established route has one or more extra hops, i.e., the number of intermediate nodes in the chain topology does not match the number of hops in the path. In conclusion, BATMAN algorithm seems to be less accurate than OLSRs

when calculating the shortest route to the destination, hence the delay and the PLR increase with the consequent impact on the estimated QoE (MOS). As an example, Fig. 4.10 depicts a significant high delay in scenarios with 50 nodes using BATMAN. These results are for the 125 m x 125 m scenario and the Free Space propagation model. Whereas BATMAN shows delays up to 1050 ms, observe that in the other cases (OLSR with 25 or 50 nodes and BATMAN with 25 nodes), the delay ranges between 20 ms and 110 ms, therefore below the ITU-T established delay limit of 150 ms for VoIP valid calls.

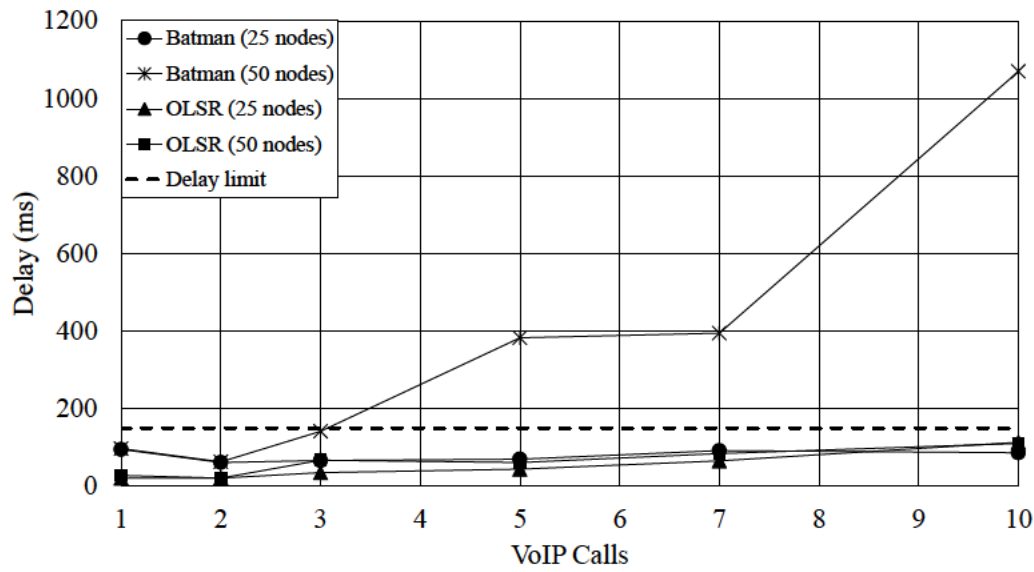


Figure 4.10. Comparison between the measured delay using BATMAN and OLSR for different number of VoIP calls and node densities in the 125 m x 125 m scenario using the Free Space propagation model.

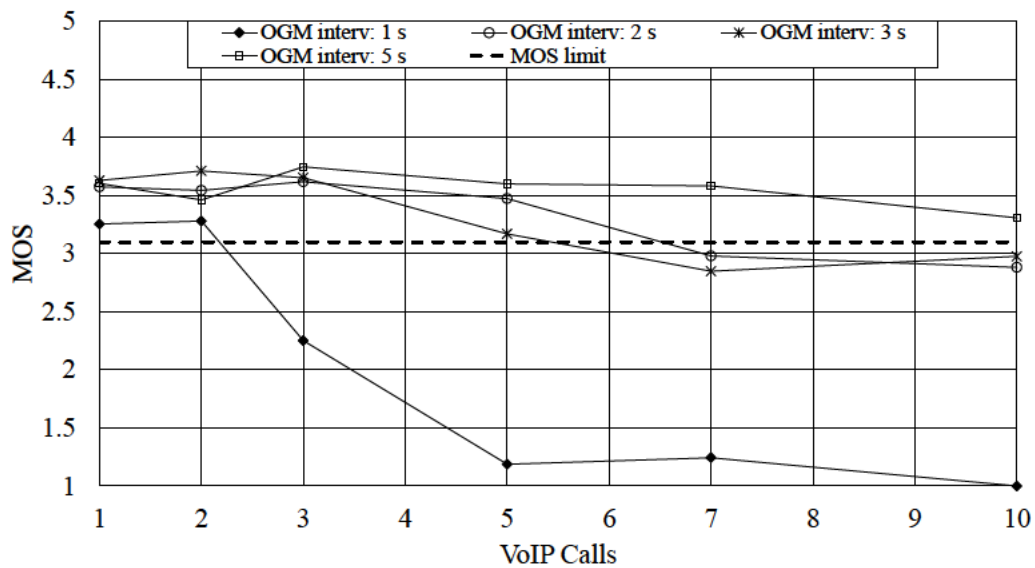


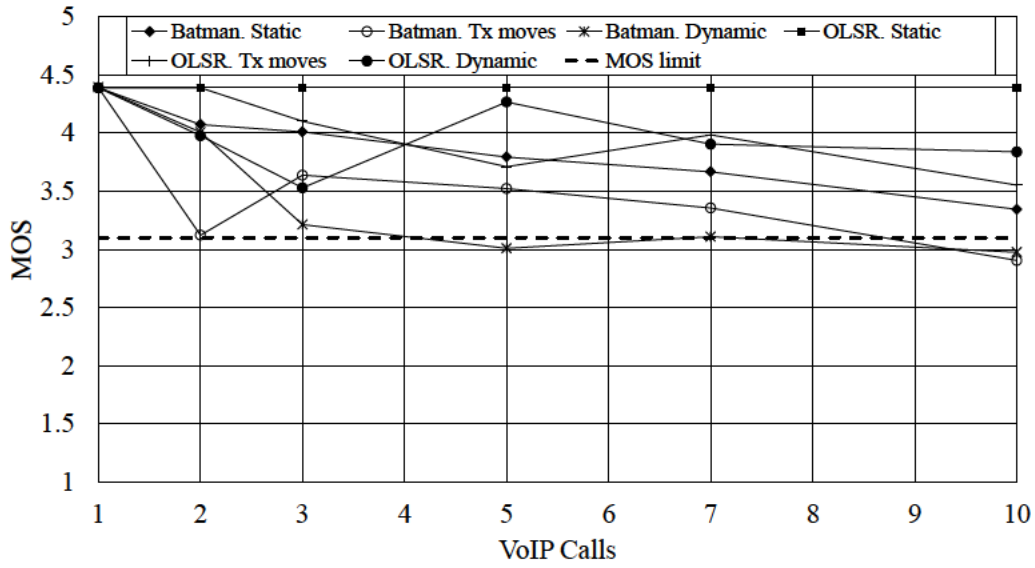
Figure 4.11. QoE (MOS) evolution for different OGM intervals and variable number of simultaneous VoIP calls in a 50 m x 50 m Free Space scenario with 50 static nodes.

Similarly to the chain scenario, the effect of the OGM interval on the performance of BATMAN affecting the VoIP QoE (MOS) is also investigated. Fig. 4.11 represents a comparison among the QoE (MOS) reached by setting different OGM intervals in a 50 m x 50 m Free Space scenario with 50 nodes. It is clear that using an interval of 5 s, the QoE (MOS) attained in every simulated scenario is greater than those obtained employing the rest of intervals. The poorest QoE performance in terms of MOS is achieved when the OGM interval takes the default value of 1 s. According to these results, BATMAN shows again the need of tuning the OGM interval depending on the network size and nodes density, in order to avoid the network collapse due to the great amount of control packets.

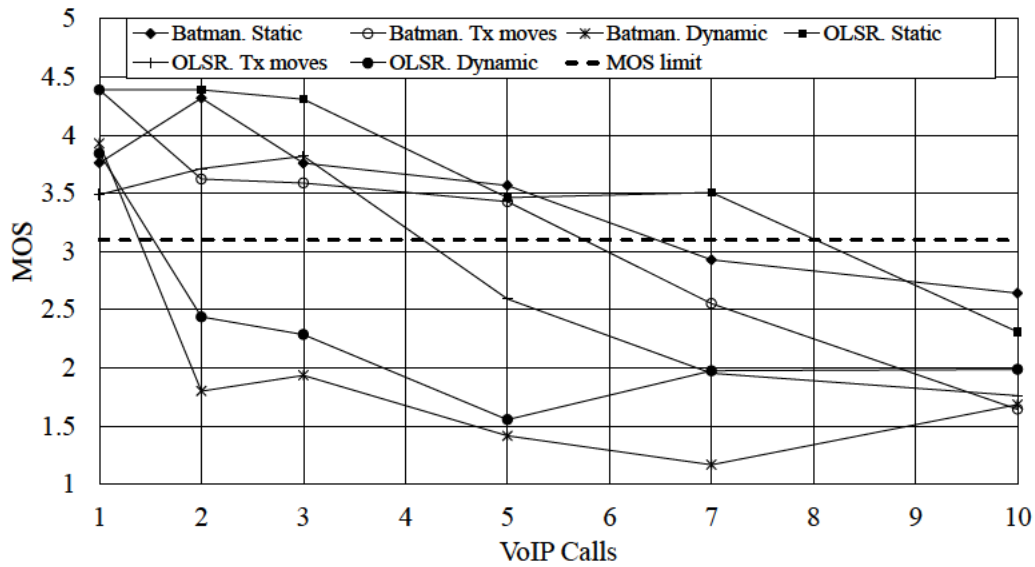
4.4.1.3 Impact of Nodes Mobility

The effect of the nodes mobility on the VoIP QoE has also been studied. Setting the mesh network topology to 25 or 50 nodes, it has been simulated three different movement patterns for the nodes: (i) all the nodes maintain their initial random position (static case, only for comparison purposes), (ii) only the VoIP TXs move (partial dynamic), and (iii) all the nodes in the network move (full dynamic) as follows. The simulated movement imitates the motion of a pedestrian handling a mobile handset device, i.e., a linear trajectory with a speed characterized by a Gaussian distribution with a mean value of 1.34 m/s and standard deviation of 0.26 m/s.

Fig. 4.12 shows a comparison between the three motion patterns under study in the 50 m x 50 m and 125 m x 125 m scenarios, for the Free Space model, and with 25 nodes. Observe that for the smaller surface scenario (Fig. 4.12 (a)), node motion has a slight impact on the quality of the calls. Both BATMAN and OLSR attain the maximum QoE (MOS) for the static configuration. In full-motion (dynamic) scenario, a little decrease in the QoE (MOS) is observed for OLSR, but always reaching QoE (MOS) values over the established minimum threshold of 3.1. However, the VoIP QoE (MOS) employing BATMAN suffers a noticeable drop under the full-motion condition. In the larger scenario (Fig. 4.12 (b)), the drop in the QoE (MOS) due to full-motion is even higher, being perceptible no matter the employed routing algorithm. Comparatively, BATMAN seems to be slightly more affected than OLSR by the nodes motion; observe that the QoE (MOS) level for a BATMAN-managed full-motion scenario is always under the QoE (MOS) estimation of the rest of the simulated conditions. The difference



(a) 50 m x 50 m

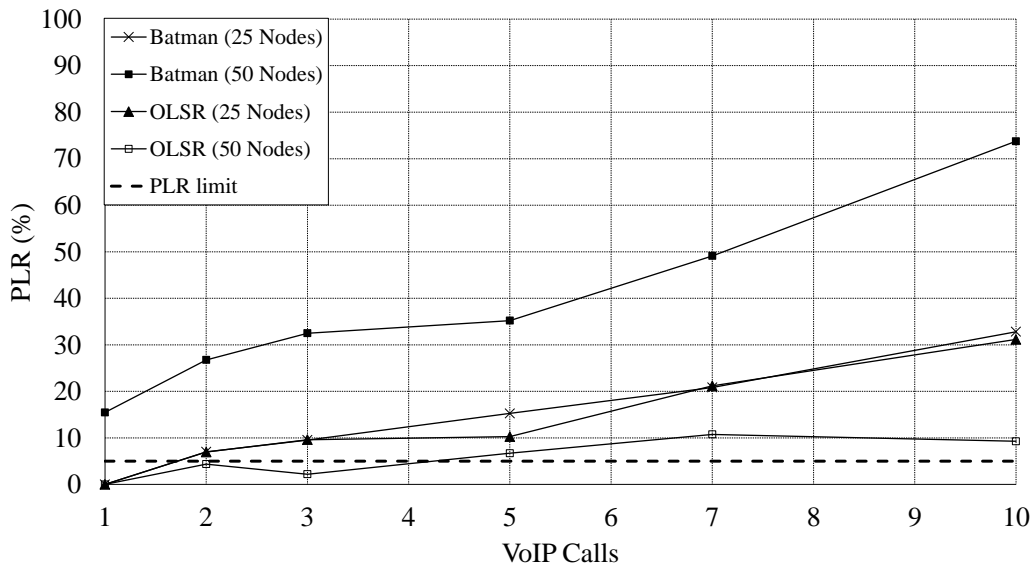


(b) 125 m x 125 m

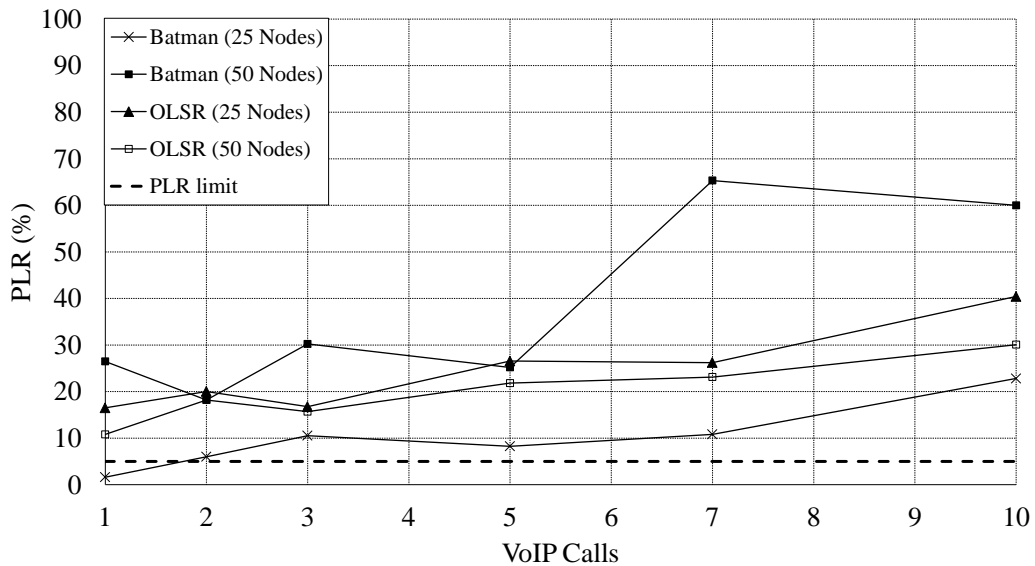
Figure 4.12. Comparison between the obtained QoE (MOS) for different movement scenarios with 25 nodes and using the Free Space propagation model.

in the results obtained for the two scenarios (50 m x 50 m and 125 m x 125 m) can be explained as follows. In the 50 m x 50 m scenario, the nodes motion does not affect the quality of the transmission since all the area is perfectly covered and there are no isolated nodes. However, in the 125 m x 125 m scenario and depending on the trajectory taken by each node, some TXs or RXs are temporarily isolated from the rest of the network due to the short coverage range provided by the low transmission power employed; for that reason a decreasing in the QoE (MOS) level of the VoIP calls is attained.

In order to clarify these outcomes from a QoS perspective, Fig. 4.13 shows the PLR suffered by the VoIP streams in the 125 m x 125 m full-motion scenario. Notice the great amount of packet losses in a 50-node Free Space scenario managed by BATMAN (Fig. 4.13 (a)). This particular scenario exhibits the main flaws of BATMAN analyzed previously. First, it illustrates the network congestion due to the OGMs-flooding, affecting especially to low loss environments as those exemplified by the Free Space model. Second, it presents the scalability problems showed by BATMAN for networks consisting of a high number of nodes, in addition to the performance degradation



(a) Free Space



(b) Nakagami-m

Figure 4.13. PLR suffered by the VoIP streams in 125 m x 125 m full-motion scenario using BATMAN and OLSR for different number of VoIP calls using the Free Space (a) and Nakagami-m (b) propagation models.

obtained in scenarios with mobility. In scenarios with low number of nodes, 25, BATMAN obtains a noticeable performance improvement. However, in this scenario with low loss, OLSR still shows a better response and the PLR suffered by the VoIP streams is clearly lower than that attained employing BATMAN. This situation turns in high-loss scenarios (see Fig. 4.13 (b)). Now, the low performance of OLSR in fading environments is manifested, and the PLR rises notably for both densities of nodes. On the other hand, BATMAN shows its robustness to fading scenarios, achieving good level of PLR for the 25 nodes scenario, but maintains its scalability issues in the network composed by 50 nodes with full motion.

As in previous sections, and in order to evaluate the impact of the BATMAN OGM interval on mobility scenarios, it has been also conducted a set of simulations in a 125 m x 125 m Free Space scenario, studying the performance of a network composed by 25 nodes. This is the same configuration used to obtain the discussed mobility results by employing the OGM by default interval (Fig. 4.12 (b) and Fig. 4.13 (a)). Thus, Fig. 4.14 depicts the QoE (MOS) accomplished for a variable number of simultaneous calls in the network and different time interval between OGMs. Observe that quite similar results are achieved, no matter the OGM interval employed. Just one call is accepted when using an interval of 1, 2, or 3 s, meanwhile, no calls are accepted as valid when tuning the interval to 0.75 or 5 s. In this scenario, the low quality of the calls is not caused by an excess of OGMs in the network, since the density of nodes is quite low. As discussed above, the VoIP quality drop is attributed to the greater size of the scenario in

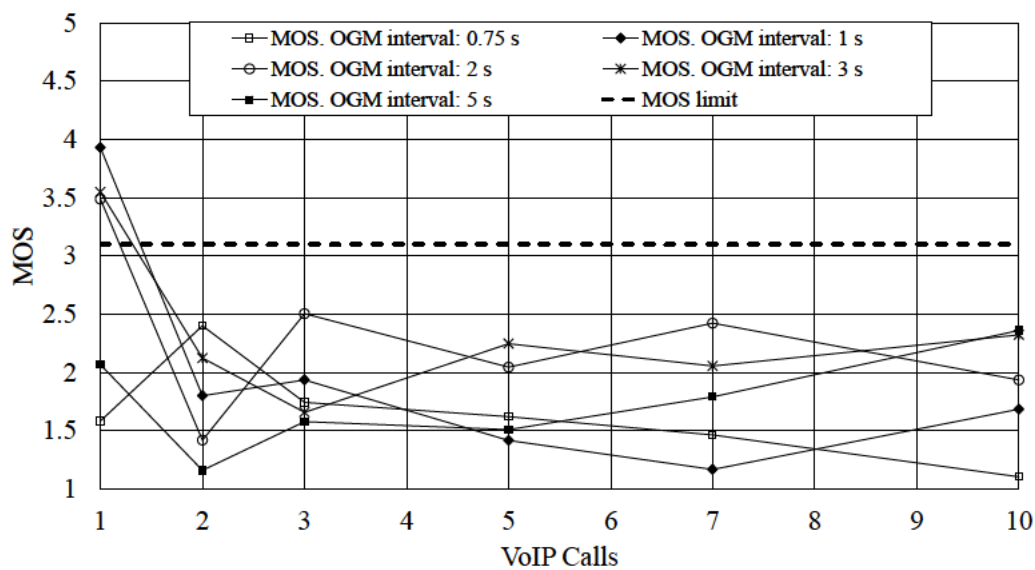


Figure 4.14. QoE (MOS) evolution for different OGM intervals and a variable number of simultaneous VoIP calls in a 125 m x 125 m Free Space scenario with 25 dynamic nodes.

comparison with the nodes coverage range, which provokes the temporary isolation of TXs or RXs during the communication. As explained in [228], currently in-developing-techniques as “waiting for destination” or “follow the destination”, not implemented yet in the current versions of the routing protocols under study, are needed to overcome this transient connectivity problems and no actions can be taken to improve the QoE (MOS) of the VoIP calls when the nodes follow “the marginal mobility model” under the management of BATMAN or OLSR.

4.4.2 Video-streaming Traffic

In this part the focus is on evaluating the capability of BATMAN and OLSR to support more demanding traffic, such as video-streaming applications. As reference results, it has been also included the outcomes obtained for a VoIP service. Moreover, it is also investigated the correlation between the results attained in two different-nature test-beds. To this end, the same scenario has been reproduced in the network simulator and in the realistic test-bench. It was configured an ad-hoc mesh network composed by 15 nodes and its performance was evaluated with a variable number (1, 2, 3, 5, and 7) of simultaneous VoIP calls and video-streams flowing through it. The floor in which the Emulab test-bench is deployed has the following dimensions: 110 m x 65 m, and it has been recreated in the Omnet++ workbench as well (see Fig. 4.15).

In the realistic scenario, five different random groups of transmitters (TX) and receivers (RX) were employed, establishing three independent communications between

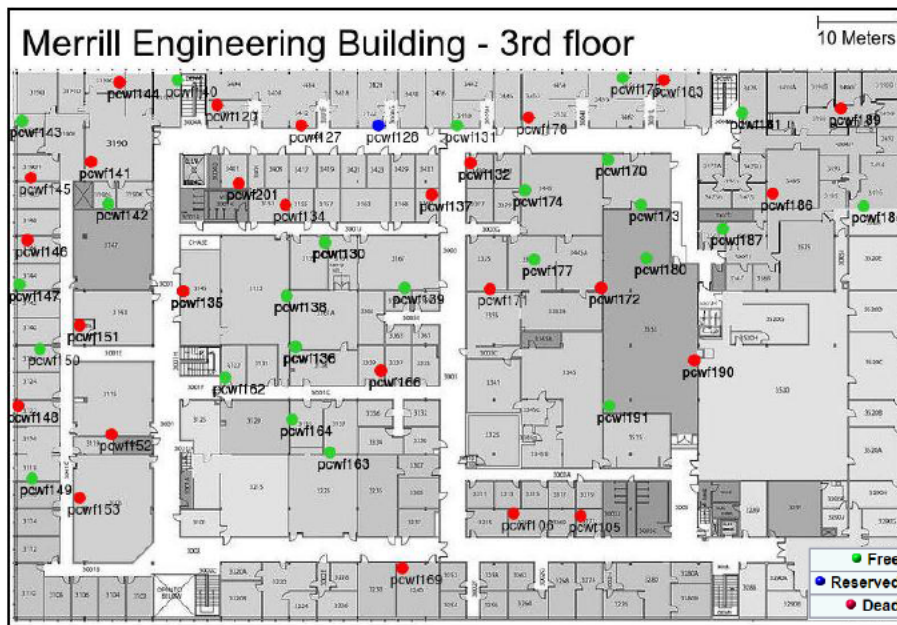


Figure 4.15. Example of Emulab test-bench map (extracted from [234]).

TABLE 4.2. Emulab nodes characteristics

Feature	Value
Emulab PC type	pc2400w
Processor	Core 2 Duo E6600
Processor speed	2.4 GHz
RAM	2048 MB
Wireless Card	Netgear WAG311 802.11a/b/g (Atheros)
Operative System	Ubuntu 12.04. 64 bits
Linux Kernel version	3.2.0
BATMAN version	Batman-adv 2014.2.0
OLSR version	olsrd 0.6.6.2

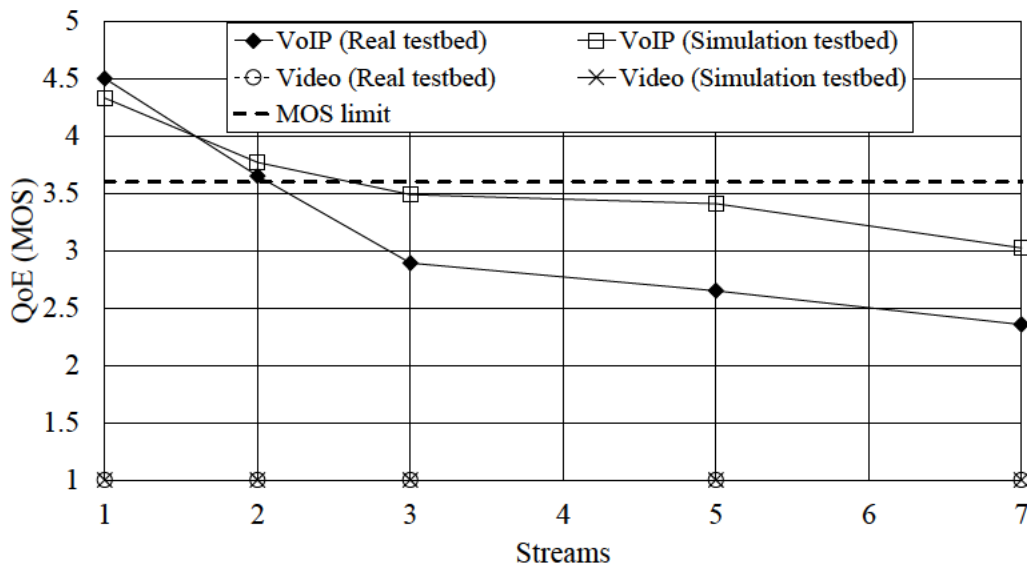
each TX-RX pair. In the simulation environment, as well as in the previous experimental study, 10 independent simulation instances with different seeds have been run for every evaluated scenario. Thereby, the results have been obtained by taking the average value for every parameter under consideration. The simulator implementations of both 802.11 and routing protocols have been configured as in the previous part (see Table 4.1). In turn, the Emulab node characteristics and the employed routing protocols versions are specified in Table 4.2.

Similar to the previous section, and as described in the ITU-T Rec. G.114 [182] and G.1010 [16] the maximum accepted PLR for VoIP services is 5% and for video-services is just 1.5%; regarding delay, the maximum accepted value for multimedia real-time communications is 150 ms. From a QoE perspective, in this section a minimum QoE (MOS) value of 3.6 is established in order to consider a multimedia transmission as valid. Observe that this value is more restrictive than that employed so far. This has been done aiming to represent the tougher requirements that video services present comparing with VoIP services.

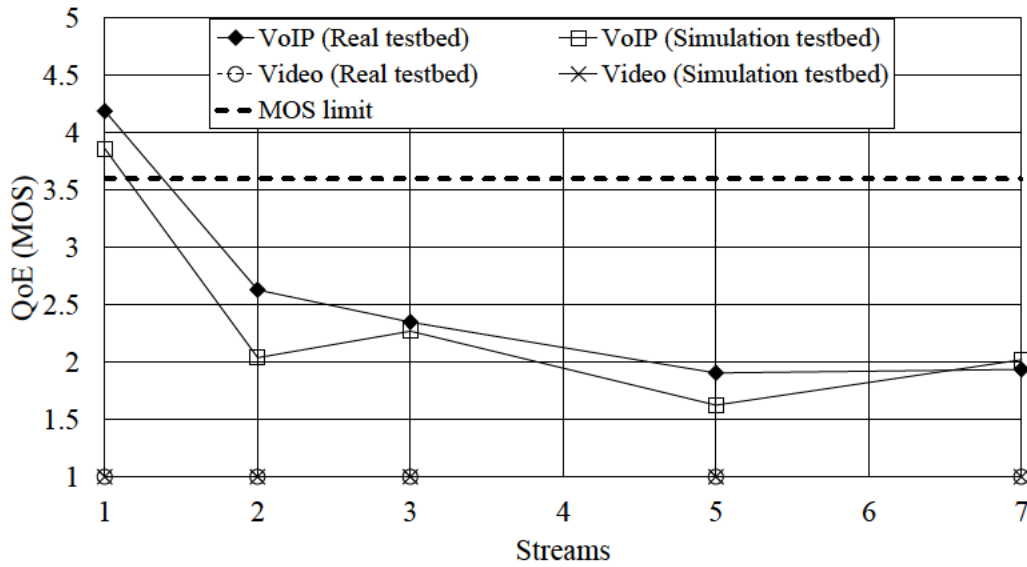
4.4.2.1 Regular Performance Evaluation

First, an evaluation of the system described above (Fig. 4.15) without any modification to the routing protocols has been performed, i.e., all protocol parameters were set to their default values as described in their corresponding RFC (please, see Table 4.1). Fig. 4.16 illustrates the QoE (MOS) attained for a variable number of simultaneous multimedia streams in the network for BATMAN and OLSR routing protocols, respectively. This figure includes the results for both test-benches, namely, the real and simulated environments. Observing the VoIP case (codec G.711), the only

acceptable QoE (MOS) values (over 3.6) were achieved for a low number of VoIP calls flowing through the network (Fig. 4.16 (a) y Fig. 4.16 (b)). Thus, using BATMAN (Fig. 4.16 (a)) the QoE (MOS) falls below the established threshold when three or more simultaneous calls flow in the network. With less number of calls, it has been obtained excellent QoE (MOS) values around 4.5 and 3.7 for 1 and 2 simultaneous calls, respectively. Note that, although the showed trend is the same for the real and the simulation measurements, the latter tends to overestimate the quality of the calls. Even so, in both test-benches the same number of simultaneous calls over the established QoE (MOS) threshold of 3.6 is obtained, namely, 2 VoIP calls.



(a) BATMAN



(b) OLSR

Figure 4.16. QoE (MOS) reached for a variable number of simultaneous multimedia streams in the network. Results for BATMAN and OLSR with VoIP and video traffic.

For OLSR (Fig. 4.16 (b)), the results are worse than those attained by BATMAN: the system just provides enough quality support for only one VoIP connection. Under OLSR, the results obtained in both test-benches are slightly closer than in the BATMAN case. However, now the real environment offers better results.

Regarding video-traffic, a Constant Bit Rate (CBR) stream of 5 Mbps (HD resolution) has been used. Note that for both protocols under study, the QoE (MOS) obtained is unacceptable in all the evaluated scenarios (Fig. 4.16 (a) and Fig. 4.16 (b)). Consequently, not even one video communication has been successfully established in the studied examples.

In order to analyze these outcomes, but from a QoS perspective, Fig. 4.17, Fig. 4.18, and Fig. 4.19 represent the delay and PLR obtained in the previously discussed scenarios. Note the high level of delay affecting the video-streams in the real test-bench no matter the routing protocol in use (Fig. 4.17). That fact causes continuous communication interruptions that provoke an important drop on the quality perceived by the users. Analogously, the minimum PLR suffered for both protocols under consideration supporting video traffic is 40%, clearly unacceptable for such highly-demanding service (Fig. 4.19 (a) and Fig. 4.19 (b)). The PLR evolution in both BATMAN and OLSR scenarios is similar and significant differences between the results obtained in the real or the simulation test-benches are not perceived. For VoIP traffic, the delay is more controlled than in the video-streaming case (Fig. 4.17). Notice

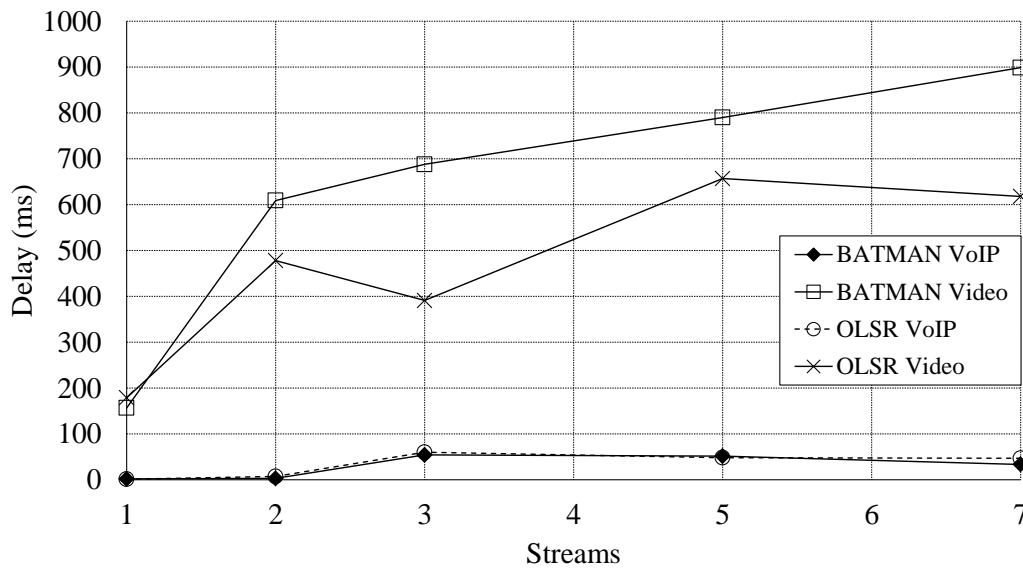
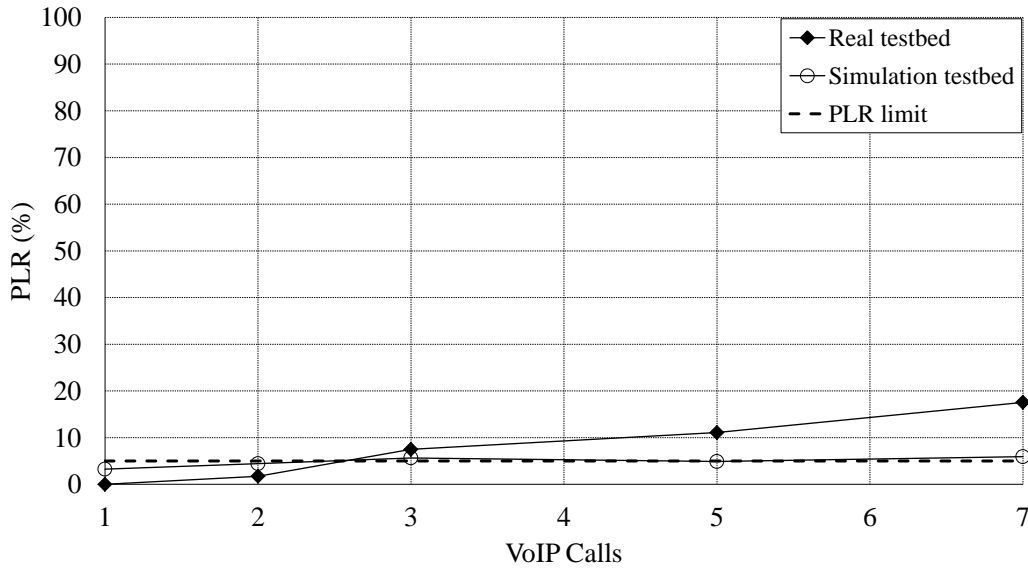
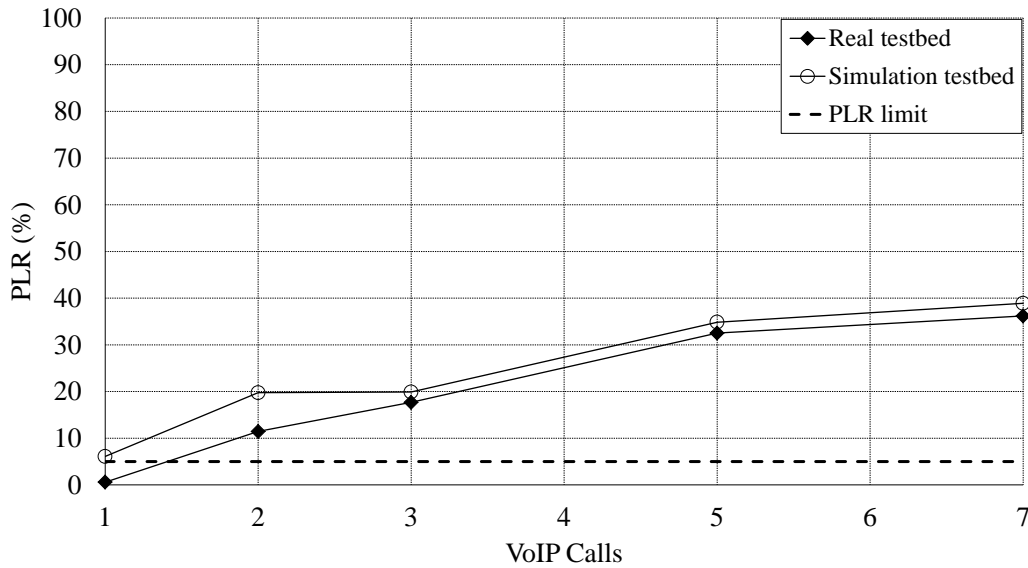


Figure 4.17. Delay evolution in the real test-bench for a variable number of simultaneous multimedia streams. Results for BATMAN and OLSR supporting VoIP and video traffic.



(a) BATMAN



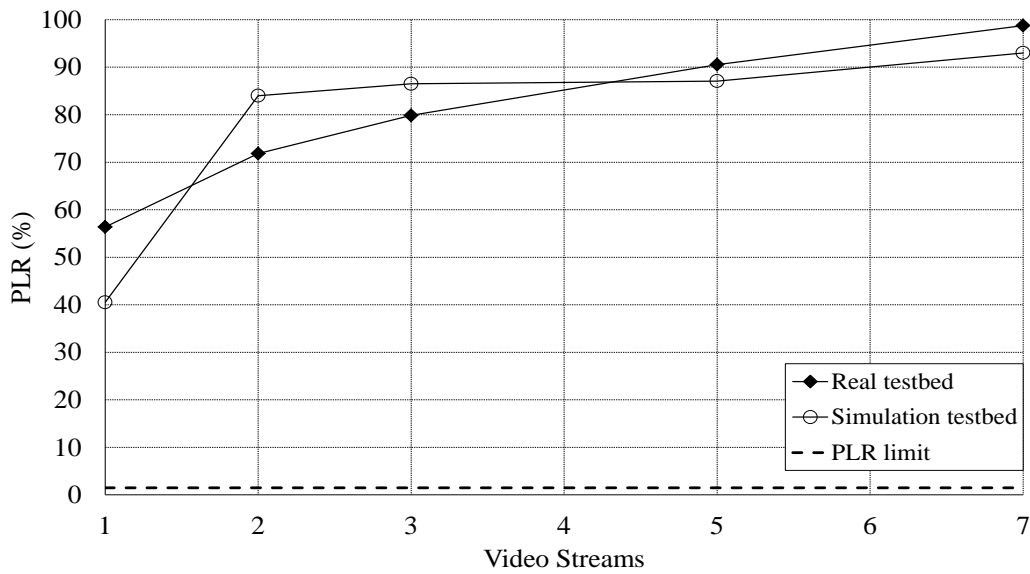
(b) OLSR

Figure 4.18. PLR attained for a variable number of simultaneous VoIP calls in the network. Results for BATMAN and OLSR.

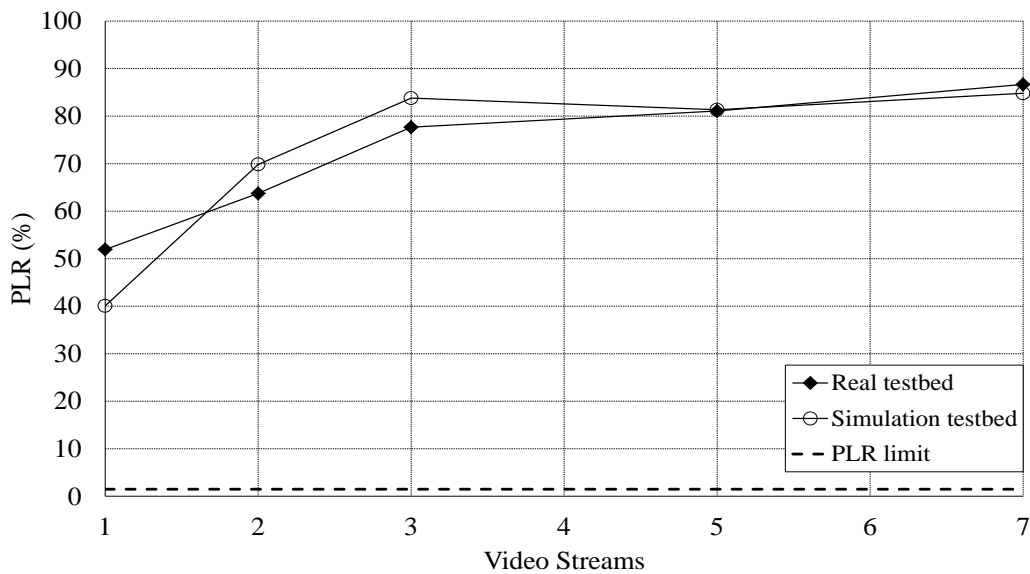
that its value is always below 80 ms, which is a quite acceptable figure. Hence, the VoIP QoE (MOS) trend showed in Fig. 4.16 (a) and Fig. 4.16 (b) is justified by the PLR suffered by the voice flows (Fig. 4.18 (a) and Fig. 4.19 (b)). Observe that just 2 simultaneous VoIP calls in the case of BATMAN and 1 VoIP call employing OLSR exhibit a PLR below the established threshold of 5%. This result is therefore consistent with the number of simultaneous calls accepted as valid from the QoE perspective discussed above (Fig. 4.16 (a) and Fig 4.16 (b)). In the light of these results, it is concluded that neither BATMAN nor OLSR in their default configurations are capable to support heavy real-time traffic such as HD video-streaming. On the other hand, they

reach acceptable quality levels just when few concurrent VoIP calls are present in the network.

With the aim of comparing both test-bench's results from a statistical perspective, a correlation analysis of the PLR obtained in both test-beds has been performed. As regards BATMAN, it has been obtained Pearson correlation coefficients (R) of 0.84 and 0.86 for VoIP and video traffic, respectively. In the case of OLSR, for the same type of traffic, it has been attained R values of 0.99 and 0.94, respectively. These statistical results confirm the high similarity between the outcomes from both test-beds.



(a) BATMAN



(b) OLSR

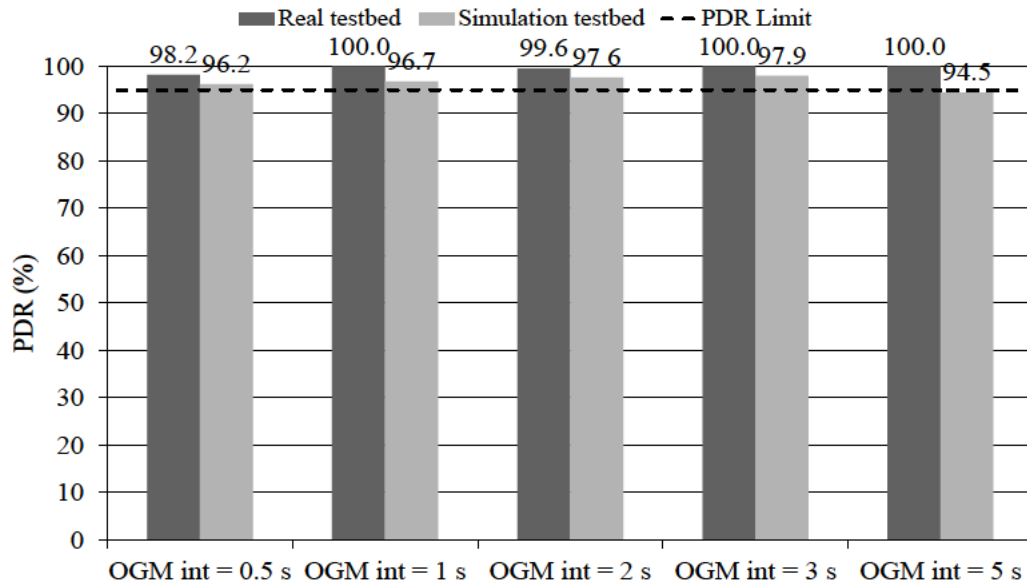
Figure 4.19. PLR attained for a variable number of simultaneous video streams in the network. Results for BATMAN and OLSR.

4.4.2.2 Tuning of Routing Messages Intervals: Impact on the VoIP Service

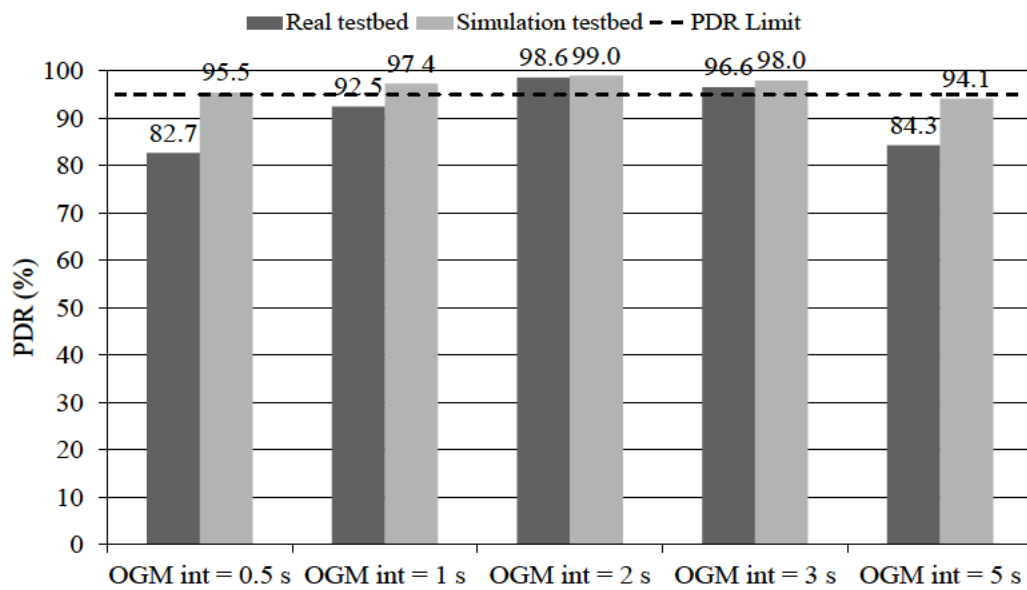
As aforementioned, the need of tuning the BATMAN OGM interval to adapt it to the actual network conditions has been suggested in the previous section. This preliminary idea is going to be tested now in a real test-bench rather than by simulation, as originally done before. In order to provide a more complete study, the OLSR HELLO interval has also been tuned. Other works [235]–[237] have previously investigated the effect of tuning the OLSR intervals but from a computer simulation assessment [235], [236] or in a real test-bench just using 5 nodes [237]. Therefore, it is included here the performance of OLSR when tuning the OLSR HELLO interval from both methodologies simulation and real tests composed by a significantly greater number of nodes.

Fig. 4.20 shows the impact of tuning the BATMAN OGM interval on the network PDR. This metric has been evaluated for a different number of simultaneous VoIP streams in the network. Observe that as the system load increases, the best results are attained for longer OGM intervals. In the case of the lowest load (1 VoIP call, Fig. 4.20 (a)), quite similar results in both test-benches are obtained, with acceptable levels of PDR reached no matter the value of the OGM intervals employed. This situation changes when more traffic is present in the network; for 3 and 5 simultaneous VoIP transmissions, Fig. 4.20 (b) and Fig. 4.20 (c) show that the best results are obtained for an OGM interval of 2 s, i.e., a larger time interval than that set by default of 1 s. Moreover, for the highest system load (7 calls, Fig. 4.20 (d)), the maximum PDR is attained with the longest OGM interval, namely, 5 s. Notice that in some scenarios, e.g., 3 or 5 simultaneous VoIP streams, a proper OGM interval tuning satisfies a PDR over the ITU-T limit of 95%, which is not possible to reach when employing the default OGM interval. Regarding the accuracy of the measurements, observe that, as in the previous subsection, the simulator tends to overestimate the BATMAN performance. However, the resulting tendency in both test-benches is the same for every scenario under evaluation.

From a user perspective, Fig. 4.21 represents the QoE (MOS) attained for 3 and 7 simultaneous VoIP calls by tuning the OGM interval. Observe how in both cases the maximum QoE (MOS) are obtained for different OGM intervals: with 3 VoIP calls in the network, the interval has to be tuned to 2 s or 3 s in order to fulfill a QoE (MOS)

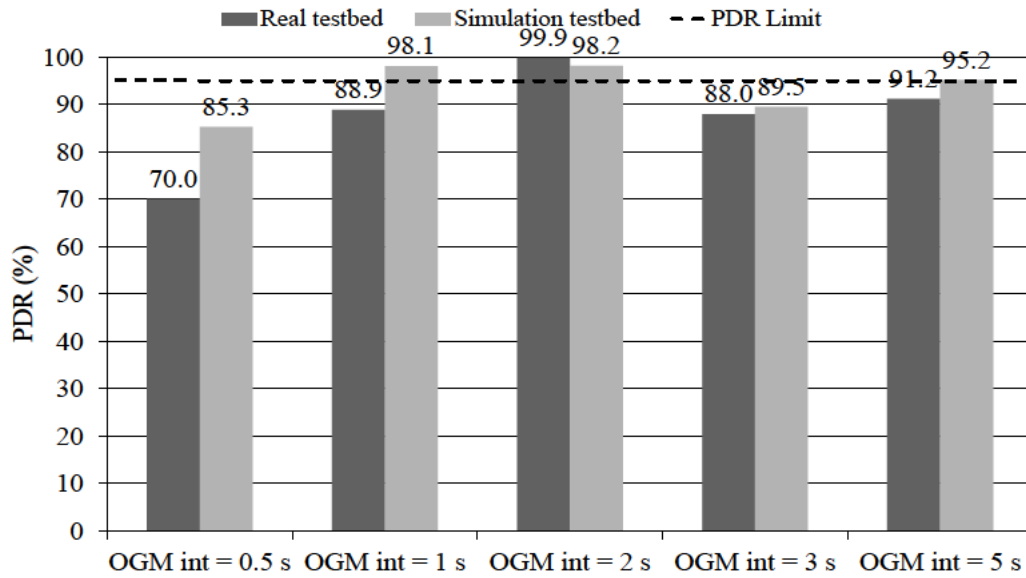


(a) 1 VoIP call

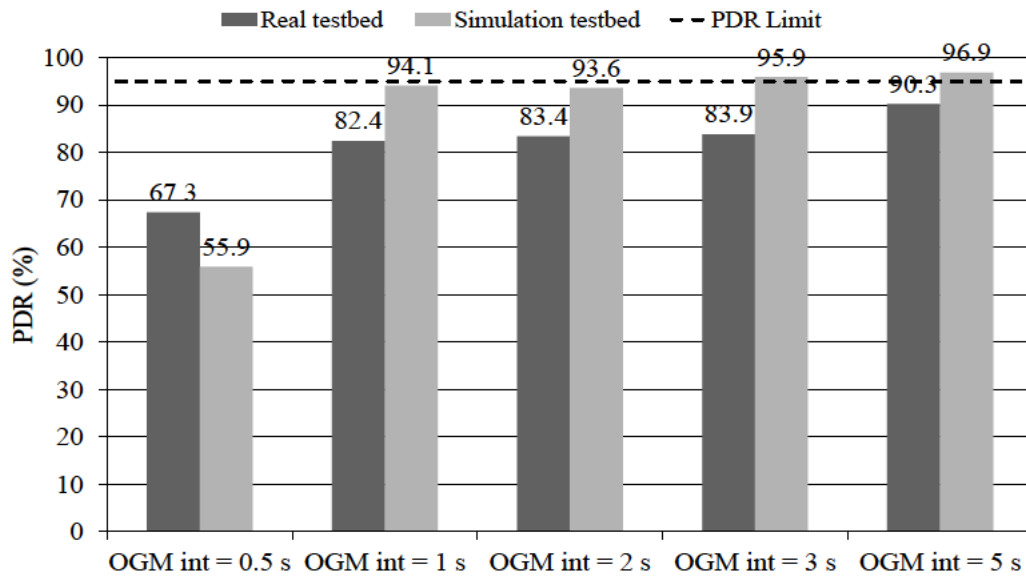


(b) 3 VoIP calls

Figure 4.20. PDR evolution for a variable number of simultaneous VoIP calls in the network, tuning the BATMAN OGM interval.



(c) 5 VoIP calls



(d) 7 VoIP calls

Figure 4.20. PDR evolution for a variable number of simultaneous VoIP calls in the network, tuning the BATMAN OGM interval.

over the established threshold of 3.6. In turn, for 7 VoIP streams in the system, the OGM interval needs to be set to 5 s to obtain the maximum quality. Nevertheless, in the last case the OGM interval tuning seems to be insufficient and the QoE (MOS) does not reach the established minimum acceptable level (QoE (MOS) = 3.6), despite of the OGM interval employed. Again, the simulation environment yields quality values above those obtained in the real test-bench. Therefore, from both QoS and QoE analyses it is deduced that, as the traffic load increases, it is advantageous to reduce the amount of control information in the network to improve the quality of the multimedia connections. This control-packets reduction is accomplished by enlarging the OGM

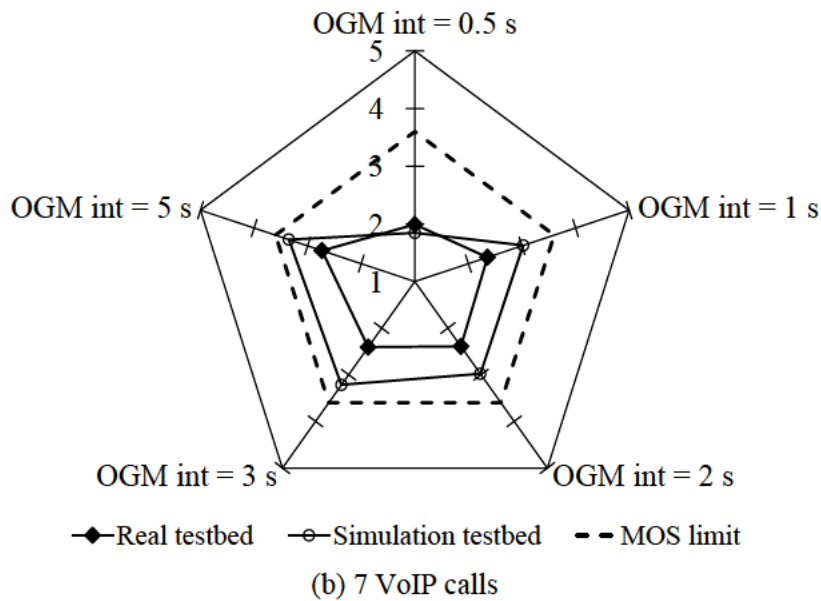
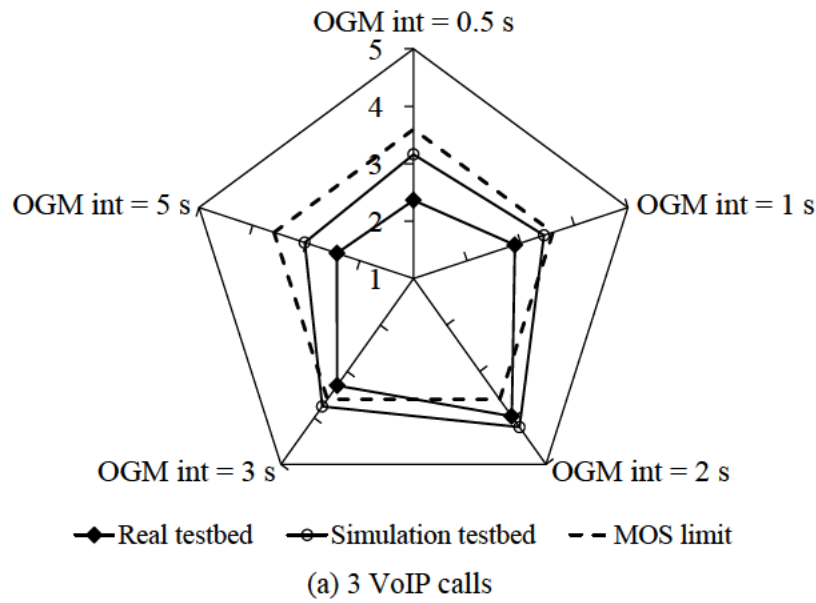
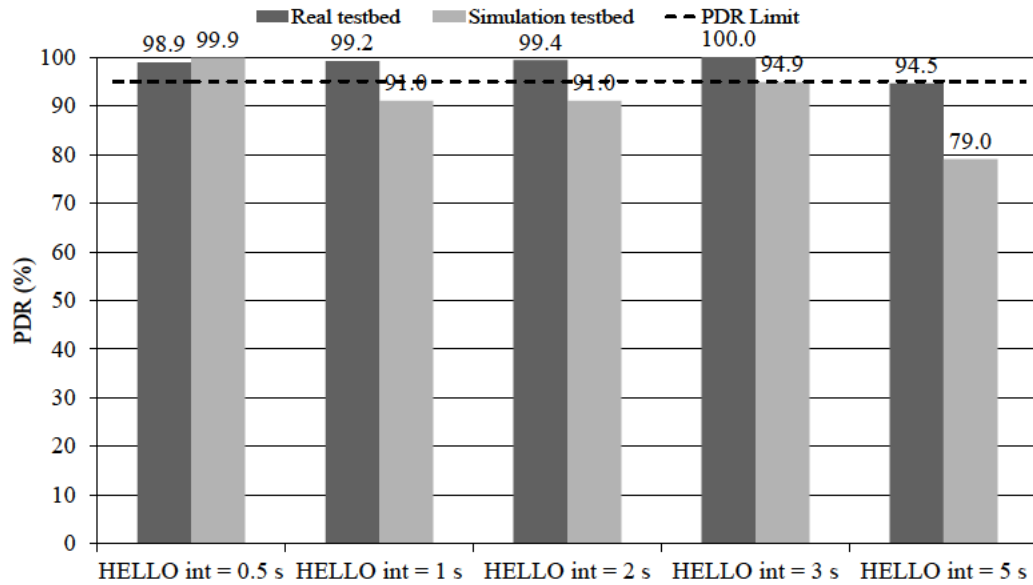


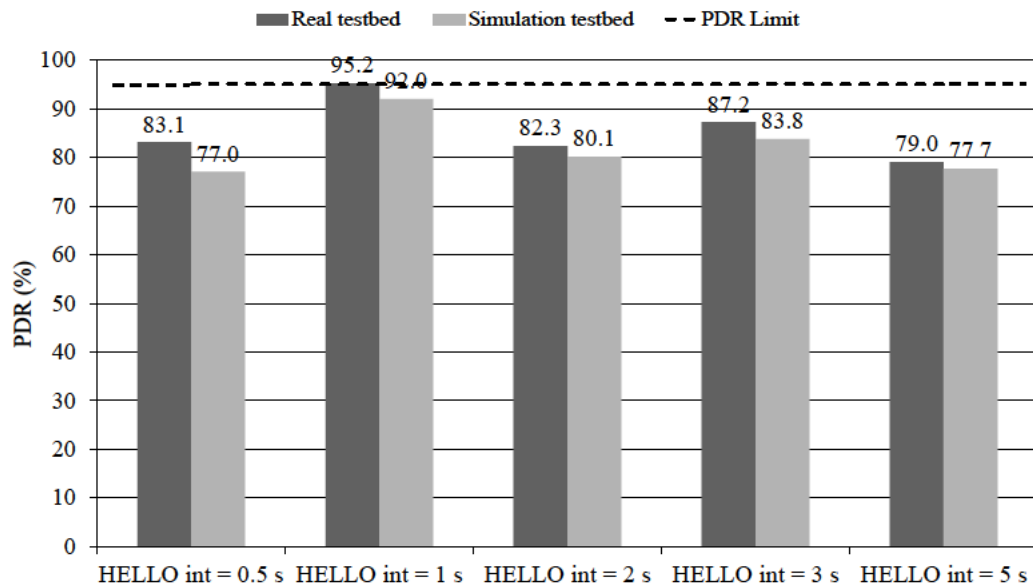
Figure 4.21. QoE in terms of MOS for 3 and 7 VoIP calls obtained by tuning the BATMAN OGM interval.

interval, which shows a great influence on the performance of the service flowing over the network.

Focusing on OLSR, it is not perceived a clear improvement in the system performance by tuning the HELLO interval. Observe in Fig. 4.22, which depicts the PDR for a variable number of simultaneous VoIP calls in the system, that the relationship between this metric and the HELLO interval is not straightforward. In [235], authors obtained greater throughput values by decreasing the HELLO interval in a network composed by nodes with high mobility. In the static scenario under evaluation, this effect is not noticeable. This outcome is consistent with those attained by Demers and Kant [236]. In this work, authors stated that, in static non-clustered networks, there is no significant change in the total traffic sent/received by the ad-hoc nodes as a function of the HELLO interval. Furthermore, this fact also agrees with the results presented by Hiyama *et al.* [237]. Although authors claimed in their conclusion that, in static scenarios, the throughput is higher as the HELLO interval increases, their results present a great variability. Furthermore, their work does not provide any statistical analysis to prove the relationship between the HELLO interval tuning and the network performance, which in the author's humble opinion detracts consistency for their statement. It is also worthwhile to remark the poor overall performance of OLSR delivering VoIP packets in comparison with BATMAN (please, compare Fig. 4.20 and Fig. 4.22); with the exception of the lowest traffic-load case (1 VoIP call, Fig. 4.22 (a)), the remaining scenarios show low PDR values, clearly inadequate to guarantee the established quality minimums. These outcomes are further confirmed from a QoE perspective (Fig. 4.23). Observe how, no matter the HELLO interval in use, the QoE (MOS) is always less than the minimum value of 3.6. With 3 simultaneous VoIP calls in the network, the best QoE (MOS) is attained by tuning the HELLO interval to 1 s. In turn, loading the system with 5 VoIP streams the best result is obtained with a HELLO interval of 3 s, although far below the established threshold of 3.6. As in the previous subsection, and on the contrary to the case of BATMAN, the simulator underestimates the OLSR performance in the majority of the evaluated scenarios.



(a) 1 VoIP call



(b) 3 VoIP calls

Figure 4.22. PDR evolution for a variable number of simultaneous VoIP calls in the network, tuning the OLSR HELLO interval.

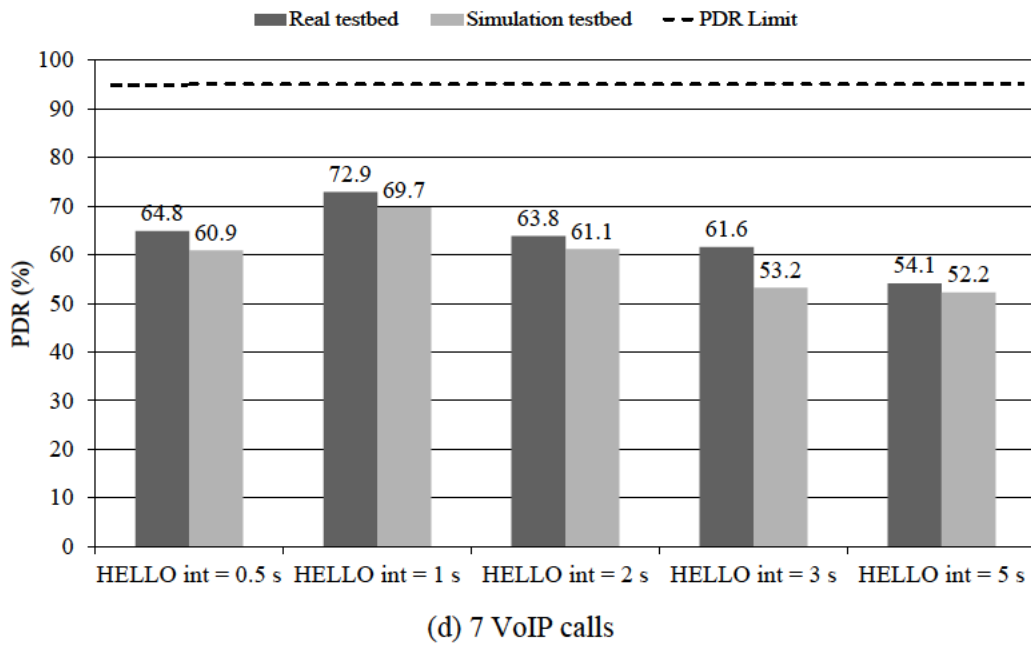
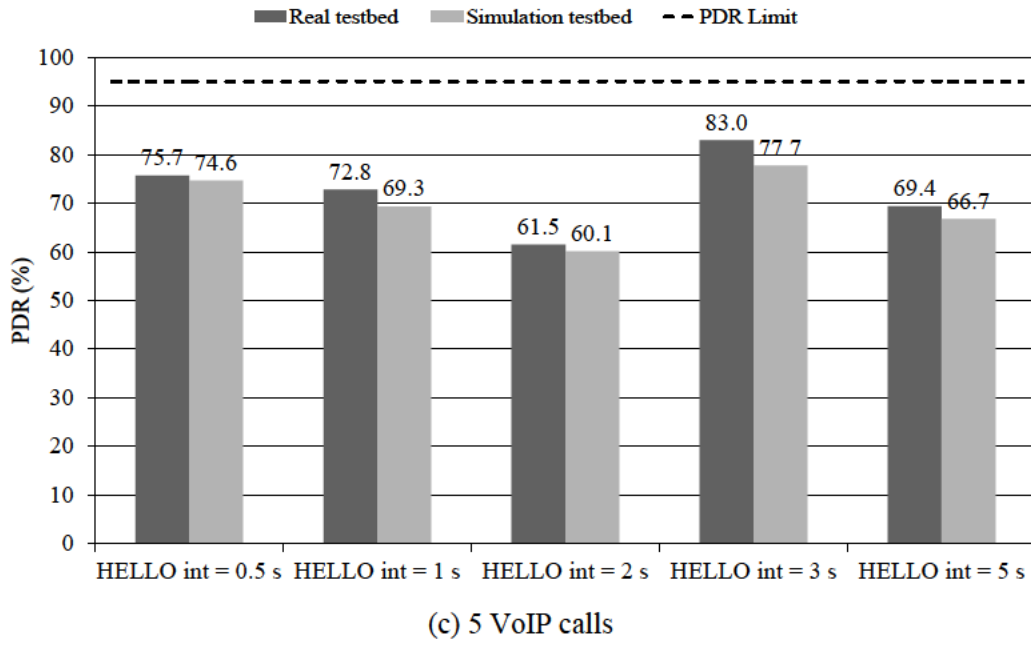


Figure 4.22. PDR evolution for a variable number of simultaneous VoIP calls in the network, tuning the OLSR HELLO interval.

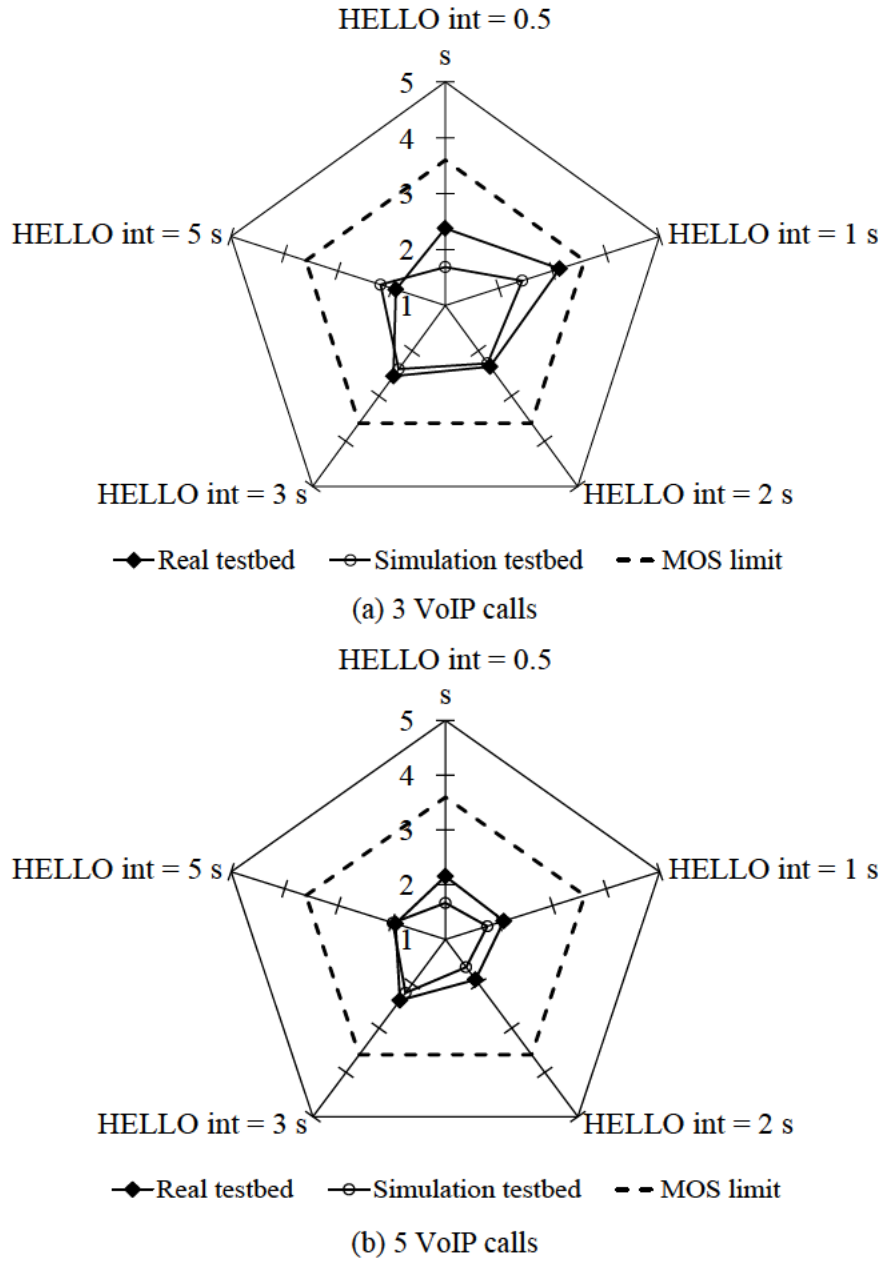


Figure 4.23. QoE in terms of MOS for 3 and 5 VoIP calls obtained by tuning the OLSR HELLO interval.

In order to provide statistical evidence of the aforementioned outcomes, a series of ANOVA (Analysis of Variance) tests at 95% level of confidence has been performed. The F statistics and p-values obtained in the ANOVA tests are shown in Table 4.3. The null hypothesis is that OGM/HELLO intervals do not have an effect on the average number of received VoIP packets. The intervals used are 0.5 s, 1 s, 2 s, 3 s, and 5 s. Results are grouped by the number of simultaneous VoIP calls in the system. First, observe that no significant differences have been obtained in any OLSR scenario, revealed by p-values higher than 0.05 (see last column of Table 4.3). On the other hand, BATMAN scenarios with low traffic-load (1 and 2 VoIP calls) seem to be neither

TABLE 4.3. One-way ANOVA results

VoIP Calls	BATMAN		OLSR	
	F	p	F	p
1	1.02	0.42	1.79	0.17
2	1.27	0.3	0.75	0.56
3	2.86	0.03	0.96	0.43
5	7.25	0.00003	1.35	0.25
7	4.31	0.002	0.51	0.73

reactive to the OGM tuning. However, for a higher traffic load (3, 5, and 7 simultaneous calls) p-values below the significance level of 0.05 are obtained. Therefore the null hypothesis (the OGM interval does not impact on the number of received VoIP packets) is rejected, and it is concluded that there is statistical evidence of differences in the VoIP packets received according to the OGM interval employed, corroborating the previous discussion. Thereby, it is determined that, in a static situation, the higher the traffic load in the network the larger the OGM interval in order to provide better quality levels. On the contrary, no clear relationship has been found between the tuning of the OLSR HELLO interval and the quality achieved by the service flowing through the network.

Regarding the fidelity of the employed test-benches (simulation and real test-bed), both of them reveal a similar behavior when tuning the routing intervals, although the simulator tends to overestimate the performance of BATMAN and to underestimate the OLSR operation. The Pearson correlation coefficient (R) attained for BATMAN and OLSR experiments corresponding to this subsection are 0.74 and 0.96, respectively.

A similar study injecting in the network the same video-traffic pattern as in previous subsections (bit-rate of 5 Mbps) has been conducted for both protocols. These results are not presented here due to the poor performance of the system that, regardless the time-intervals employed, did not reach acceptable quality levels. For that reason, the maximum video-stream bit-rate accepted by the network in order to provide this service with minimum tolerable quality levels is studied in next subsection.

4.4.2.3 Establishing the Maximum Video Bit-rate

In order to provide an overview of the network capability to support video traffic, a simulation study varying the bit-rate of the video-flows has been performed. With this

aim, the system evolution has been evaluated by changing the video-flows bit rate from 5 Mbps down to 70 Kbps. In particular, typical bit-rates matching widely-used video resolutions have been tested: 70 Kbps and 100 Kbps (QCIF), 250 Kbps and 500 Kbps (CIF), 750 Kbps and 1 Mbps (D1), and, finally, 3 Mbps and 5 Mbps (HD). Fig. 4.24 shows the PLR attained in the video-streams set to bit-rates up to 1 Mbps. Higher bit-rates corresponding to HD resolution lead the system to unacceptable packet loss levels, so it has been considered not necessary to present these results here. BATMAN (Fig. 4.24 (a) and Fig. 4.24 (c)) noticeably exhibits better performance than OLSR (Fig. 4.24 (b) and Fig. 4.24 (d)) in almost every evaluated scenario. Observe that for the former, the system accepts 1 video-stream between 500 and 700 Kbps (D1). Additionally, 3 CIF video communications of 250 Kbps and 3 QCIF streams (100 Kbps) attain acceptable levels of packet loss. With more streams in the system, the bit-rate should be decreased and, in the worst case scenario (7 video-streams), the bit-rate should be extremely reduced in order to keep the PLR down to acceptable values. On the other hand, OLSR presents rather disappointing results supporting video traffic (Fig. 4.24 (b)): PLR values below the established limit of 1.5% are obtained just for the case of one video-flow streamed at the lowest considered bit-rates (70 and 100 Kbps). In order to evaluate the fidelity of the results, please observe that the obtained PLR in both test benches (Fig. 4.24 (a) and Fig. 4.24 (b) vs. Fig. 4.24 (c) and Fig. 4.24 (d)) is quite similar. Although some particular values that differ in both test-beds are found, the Pearson correlation coefficient (R) attained for both BATMAN and OLSR tests is 0.84.

4.4.2.4 Tuning of Routing Messages Intervals: Impact on the Video-streaming Service

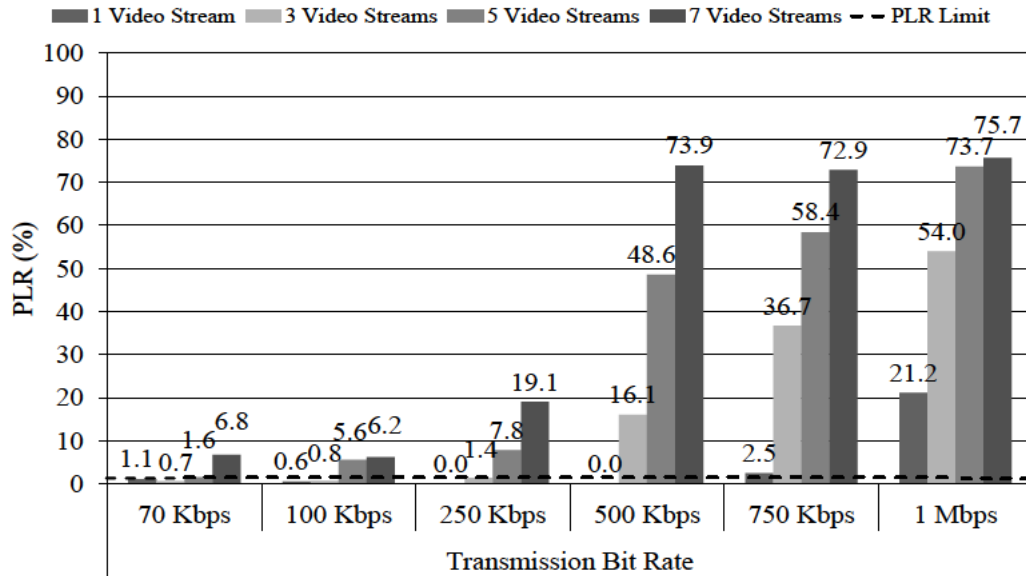
Finally, it has been conducted, for the real test-bench, a similar study to that shown for the case of VoIP traffic (Fig. 4.20), regarding the impact of a proper control-packet's time-interval tuning on the quality of the video service flowing through the ad-hoc network. Three different video-traffic patterns, namely, QCIF (70 Kbps), CIF (250 Kbps), and D1 (750 Kbps), and 5 different values for the OGM/HELLO interval (0.5 s, 1 s, 2 s, 3 s, and 5 s) have been employed. Network statistics have been collected for every scenario under study, focusing on the average PLR suffered by the receiving nodes. Table 4.4, Table 4.5, and Table 4.6 show the average PLR for each video-traffic pattern (QCIF, CIF, and D1, respectively) when the network is loaded with different number of simultaneous video-streams. BATMAN is the routing protocol tested and the

best OGM interval has been used according to network conditions. Additionally, the attained results employing the OGM interval by default (1 s) are also presented for comparison purposes.

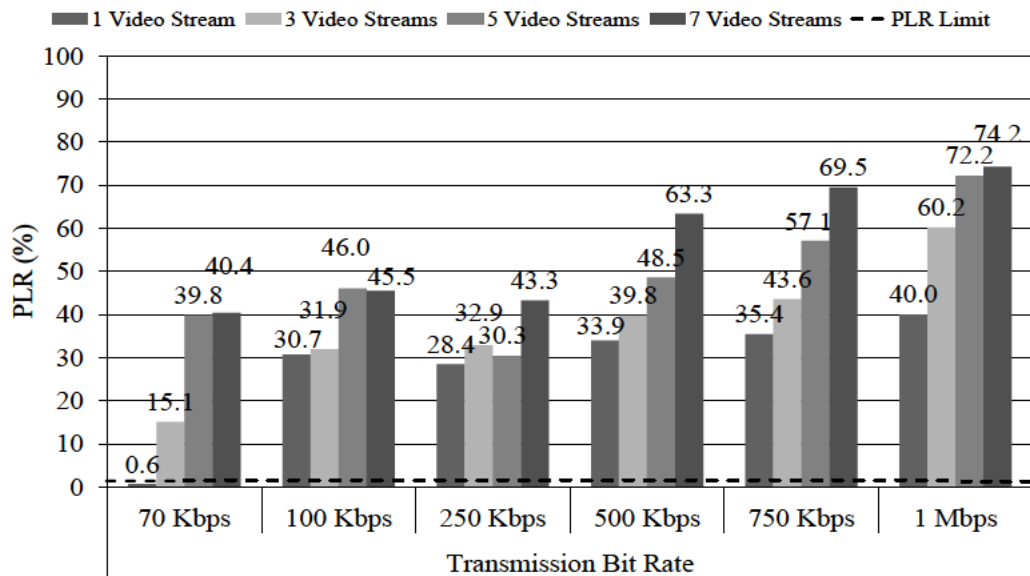
First, observe an overall improvement of the system performance, comparing with that attained with the OGM interval set by default. Furthermore, notice the evolution of the best OGM interval for each scenario: as the traffic load in the network increases, the best results are attained with greater values of the OGM interval. This confirms the statistical analysis presented for VoIP traffic (Table 4.3) from which the need of increasing the OGM interval when the network load grows was extracted. For the lightest video-traffic pattern (QCIF), the lowest PLR values are obtained by employing an OGM interval of 1 s. In turn, for heavier traffic (CIF) the best figures are attained for OGM intervals of 2 s and 3 s. Finally, for the heaviest video-traffic (D1), the best results are achieved by tuning the OGM interval to even greater values: 3 s and 5 s.

On the other hand, a similar study performed for OLSR ratifies the outcomes showed previously for the VoIP traffic case. The HELLO interval tuning does not have any impact on the quality of the multimedia service, in terms of good-put. As a sample, Table 4.7 presents the average PLR attained in the network supporting a video-streaming service of 100 Kbps (QCIF) for a variable number of flows in the network and applying the HELLO interval tuning. Observe that the poor results are quite similar no matter the HELLO interval employed: PLR figures under the established threshold of 1.5% are only obtained for just one video-stream in the network.

In view of these outcomes it is concluded that OLSR is unable to support heavy traffic services such as video-streaming, at least in its configuration by default. In turn, BATMAN has shown to be more robust managing video traffic and, even with a proper tuning, could be a real alternative to provide support to multimedia communications with good quality in MANETs.

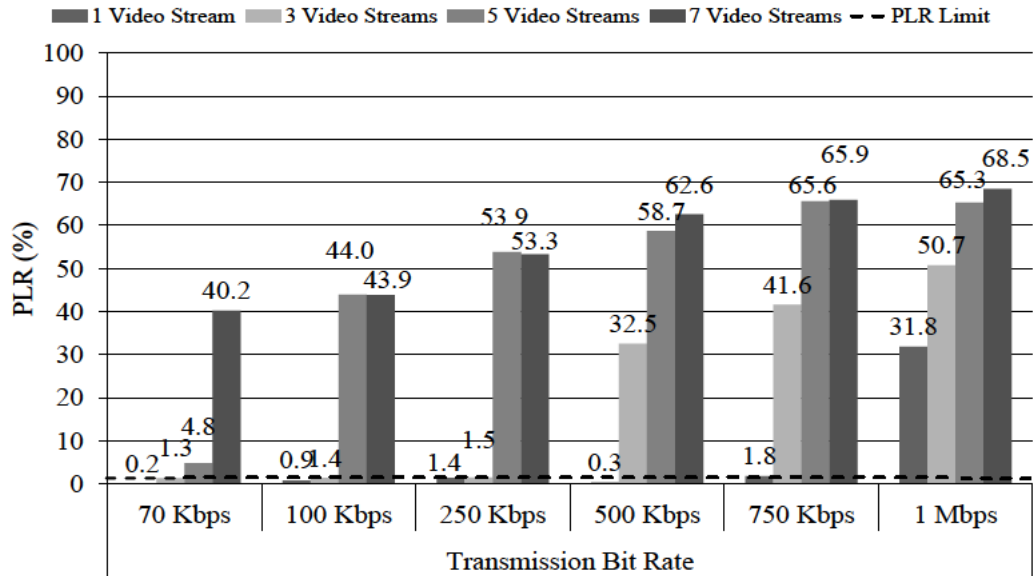


(a) BATMAN. Simulation test-bench.

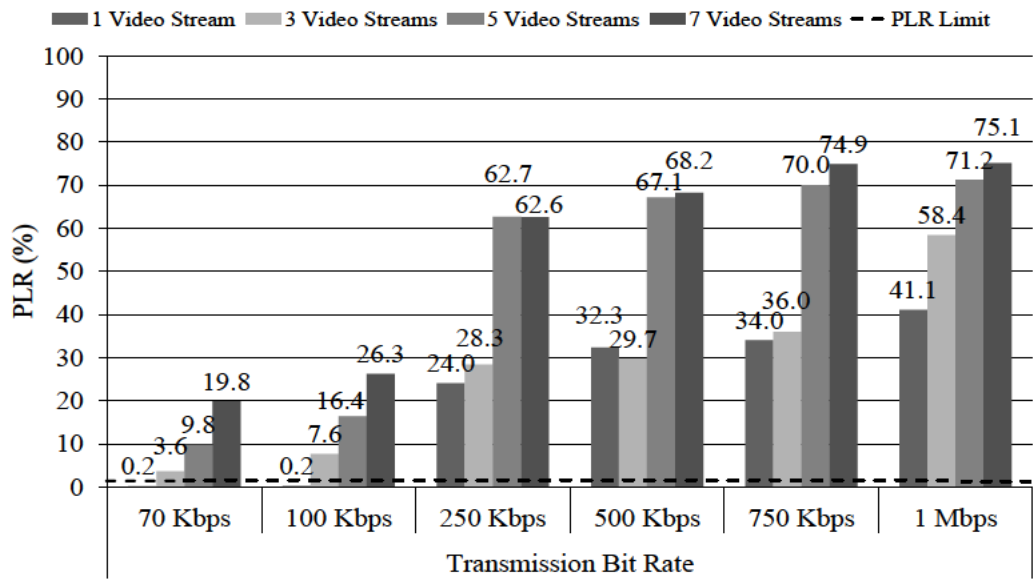


(b) OLSR. Simulation test-bench.

Figure 4.24. PLR attained for a variable number of simultaneous video-streams with different bit-rates in the network. Results for BATMAN and OLSR using their default configuration parameters.



(c) BATMAN. Real test-bench.



(d) OLSR. Real test-bench.

Figure 4.24. PLR attained for a variable number of simultaneous video-streams with different bit-rates in the network. Results for BATMAN and OLSR using their default configuration parameters.

TABLE 4.4. PLR attained for QCIF video traffic (70 Kbps) using the best OGM interval according to network conditions (results from real test- bench)

Video-streams in the Network	Best OGM Interval (s)	PLR (%)	PLR with OGM Interval = 1 s (Default Value) for Comparison Purposes (%)
1	1	0.5	0.5
3	1	1.3	1.3
5	1	1.5	1.5
7	1	13.5	13.5

TABLE 4.5. PLR attained for CIF video traffic (250 Kbps) using the best OGM interval according to network conditions (results from real test- bench)

Video-streams in the Network	Best OGM Interval (s)	PLR (%)	PLR with OGM Interval = 1 s (Default Value) for Comparison Purposes (%)
1	3	0.72	4.4
3	3	1.3	34.5
5	2	19.5	53.7
7	2	30	50.7

TABLE 4.6. PLR attained for D1 video traffic (750 Kbps) using the best OGM interval according to network conditions (results from real test- bench)

Video-streams in the Network	Best OGM Interval (s)	PLR (%)	PLR with OGM Interval = 1 s (Default Value) for Comparison Purposes (%)
1	5	0.6	11.8
3	3	29.3	35
5	3	34	55.7
7	5	60	81.9

TABLE 4.7. Average PLR attained for 100 Kbps video traffic by tuning the HELLO interval (results from real test-bench)

HELLO Interval	1 Video-stream	3 Video-streams	5 Video-streams	7 Video-streams
1s	0.46 %	9.42 %	16.89 %	26.83 %
2 s	0.19 %	7.62 %	16.38 %	26.26 %
3 s	1.94 %	15.57 %	17.13 %	24.83 %
5 s	1.3 %	19.49 %	21.91 %	23.28 %

4.5 Conclusion

In this chapter, a QoS and QoE exhaustive performance evaluation study about the capability of the ad-hoc routing protocol BATMAN to support multimedia traffic (including both VoIP and video-streaming) was conducted. The results were obtained through simulation as well as from a real test-bench, showing a high correlation between them. The achieved outcomes employing BATMAN were compared with those attained by the well-accepted routing protocol OLSR. First, the effect of the physical layer on the performance of the routing protocols was analyzed. BATMAN obtained better results than OLSR in scenarios under fading conditions characterized by the Nakagami- m propagation model, which is a hostile medium with fading channels. This behavior is related to the bigger size of OLSR control packets in comparison with BATMAN shorter control packets, which make them more exposed to the impairments introduced by the channel. On the other hand, OLSR over performed BATMAN in Free Space environments. Some problems regarding the BATMAN OGM interval were noticed; using the default value of 1 s, the OGM messages saturated the network, so the system capacity and the multimedia QoE (MOS) dropped dramatically in Free Space scenarios, since they present a lower packet loss rate. Thus, it was shown that the system performance can be improved by increasing the OGM interval, e.g., up to 2 s. In turn, OLSR showed to be unreactive to the modification of the HELLO interval. Afterwards, the impact of the number of nodes and nodes density on BATMAN performance was studied. It was detected some scalability issues in this routing algorithm operating in mesh networks having a medium/high number of nodes. Such behavior is attributed to the BATMAN path-finding algorithm, which always obtained longer routes than those calculated by OLSR. However a possible solution to these issues was also found, by tuning properly the OGM interval. Then, the effect of nodes motion on the estimated QoE (MOS) was evaluated. It was showed that, whereas its effect is hardly noticeable in small-sized networks, a QoE (MOS) drop appeared when the nodes were spread on greater areas, especially for BATMAN. Finally, it was also studied the maximum video-stream bit-rate supported by each protocol in order to provide acceptable levels of quality, again by simulation and by real experimentation. The obtained results reveal that BATMAN is capable of managing video traffic with an adjustment of its bit-rate according to the number of streams in the network. In contrast,

OLSR showed a poorer performance supporting video-streaming and the quality attained for this service seemed to be clearly insufficient.

The main conclusion derived from this chapter is the superior performance of BATMAN supporting demanding traffic in comparison with OLSR. Furthermore, it was revealed that with an appropriate tuning of BATMAN configuration-intervals, its performance can be additionally improved. In the next chapter, this idea is further investigated and a new ad-hoc routing protocol is proposed with the aim of overcome the existing traditional ad-hoc routing protocols by means of considering a novel routing paradigm: opportunistic networks.

4.6 Contributions Related to this Chapter

R. Sanchez-Iborra, M.-D. Cano, and J. Garcia-Haro, “Performance evaluation of BATMAN routing protocol for VoIP services: a QoE perspective,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4947 – 4958, 2014. (I. F.: 2.496).

R. Sanchez-Iborra, M.-D. Cano, J. J. P. C. Rodrigues, and J. Garcia-Haro, “An experimental QoE performance study for the efficient transmission of high demanding traffic over an ad-hoc network using BATMAN”, *Mobile Information Systems*. In press. (I. F.: 0.949).

R. Sanchez-Iborra and M.-D. Cano, “An approach to a cross layer-based QoE improvement for MANET routing protocols,” *Network Protocols and Algorithms*, vol. 6, no. 3, pp. 18 – 34, Aug. 2014.

R. Sanchez-Iborra and M.-D. Cano, “QoE-based performance evaluation of video transmission using the BATMAN routing protocol,” in *17th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM 2014) - 10th ACM International Symposium on QoS and Security for Wireless and Mobile Networks (Q2SWinet 2014)*, 2014, pp. 9 – 16.

R. Sanchez-Iborra and M.-D. Cano, “Estudio de las prestaciones del protocolo de enrutamiento BATMAN con tráfico VoIP,” in *XI Jornadas de Ingenieria Telematica (JITEL 2013)*, 2013, pp. 163 – 168.

R. Sanchez-Iborra and M.-D. Cano, “Laboratorios en la nube: un caso práctico,” *Anu. Jóvenes Investig.*, vol. 8, no. 1, pp. 54 – 56, 2015.

Chapter 5

Novel Proposals for Improving QoE and Energy Efficiency in 802.11g Mobile Ad-hoc Networks

5.1 Introduction

In the previous chapter, it was showed the limited capability of traditional ad-hoc routing protocols to support highly demanding services, such as multimedia streaming. It was also presented some enhancements to a particular protocol (BATMAN), which permitted increasing the quality of these services. In this chapter, the performance improvement of ad-hoc networks is further explored. Concretely, the focus is on a novel paradigm that is showing promising results regarding multimedia content distribution in ad-hoc networks: opportunistic routing. Opportunistic networks have emerged as a new networking-paradigm that is attracting the research community's interest due to its potential for enhancing communications between mobile smart devices [238]. These networks are an evolution of the Mobile Ad-hoc NETworks (MANETs), including new functionalities that make them more efficient than their precursor. Specifically, opportunistic networks take advantage of the broadcast nature of the wireless networks, i.e., direct communications between two nodes can be overheard by nearest neighbors. In ad-hoc multi-hop networks, traditional routing protocols such as OLSR (Optimized Link State Routing) [207], AODV (Ad-hoc On-Demand Distance Vector) [209], or BATMAN (Better Approach To Mobile Ad-hoc Networking) [204] calculate a unique route between transmitter and receiver. Thus, each node just considers one single neighbor as the next hop to reach a given destination. However, with opportunistic routing protocols each node selects a set of its neighbors, referred to as candidates, as the potential next hops towards the final destination. The way each node selects its candidates and how they coordinate each other to pick the most proper candidate as the actual forwarder are the two key challenges in opportunistic routing. These characteristics determine the effectiveness of the opportunistic routing algorithm [239].

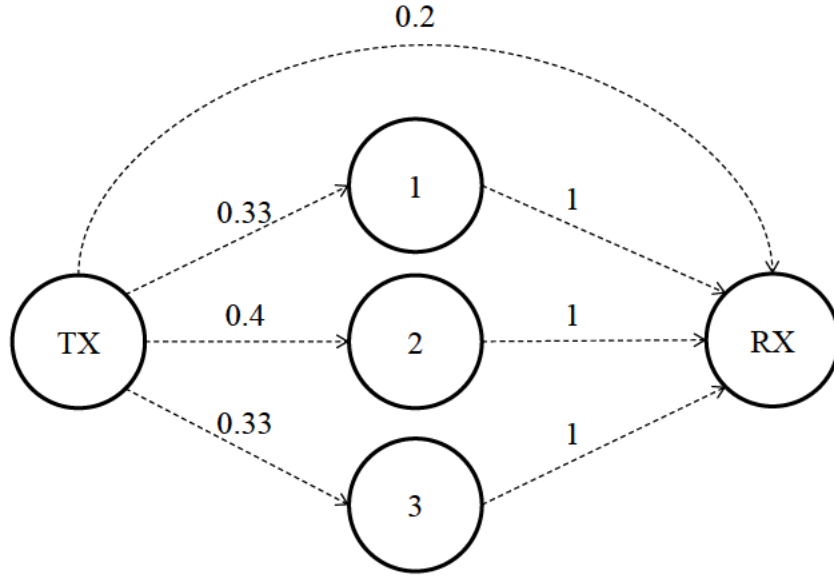


Figure 5.1. Example network.

As an example, consider the communication between the transmitter (TX) and the receiver (RX) in the network illustrated in Fig. 5.1, which also presents the delivery probability of each link. A traditional ad-hoc routing protocol that considers this metric to calculate routes, would compute the following path $\text{TX} \rightarrow \text{node 2} \rightarrow \text{RX}$ to establish the communication between the end nodes. Therefore, it would be needed 3.5 transmissions on average to reach the RX node from the TX node: $1/0.4 = 2.5$ tries to transmit a packet from TX to node 2, and an additional transmission to reach RX from the forwarder. This way of calculating routes is inefficient as explained as follows:

1) By employing an opportunistic protocol that considers the three intermediate nodes (1, 2, and 3) as candidates to reach the RX from the TX, it would be generated a “virtual link” with a delivery ratio of 0.73 ($1 - (1 - 0.4)(1 - 0.33)(1 - 0.33) \approx 0.73$); thus, only 2.37 transmissions would be necessary on average to communicate both end points (RX and TX): $1/0.73 \approx 1.37$ tries to transmit a packet from TX node to the candidates, and another transmission to reach the final destination.

2) 1 out of 5 packets directly reaches the RX node ($1/0.2 = 5$). This phenomenon is known as lucky long transmission. Traditional protocols do not take advantage of it, so these valid packets would be discarded in the RX node, which also has to wait for the intermediate node 2 to forward the packet. Besides the wrong decision of dropping a useful packet, the communication would also suffer a longer delay due to the processing time added by the forwarding node. By using an opportunistic protocol, the receiver would take this packet as valid improving the communication performance.

3) In case of a link drop, due to link failures or nodes motion, traditional protocols have to recalculate the route after noticing these troubles, which consumes extra time, adding more delay to the communication. By employing opportunistic protocols, at least one backup node would be ready to automatically forward the packet if the primary forwarder fails. For these reasons, opportunistic routing is considered as one of the best-suited routing techniques for future wireless mobile networks.

Another important point affecting ad-hoc networks is the terminals' energy consumption. Unlike the electronic technology, batteries' capacity does not follow the Moore's law [240], so extending end-devices lifetime is a greatly valued asset. Wireless cards energy consumption has a remarkable weight in mobile devices' power draining [241], thereby, it is necessary the development of networking protocols and procedures as efficient as possible in order to extend battery duration. Given the higher efficiency of opportunistic protocols compared to traditional proposals, the former may represent a real alternative for reducing the energy consumed in routing tasks.

In this chapter, a novel opportunistic routing algorithm is presented. This protocol, so called JOKER (auto-adJustable Opportunistic acKnowledgment/timEr-based Routing), gets some functional features from the pro-active ad-hoc routing protocol BATMAN [204]. The main basis taken from BATMAN is its simplicity regarding routing tasks, which entails low computational and memory needs in the nodes, making it suitable for mobile devices with limited processing power. In the previous chapter, it was showed the capability of BATMAN to support multimedia traffic in low-consumption nodes by tweaking some of its configuration parameters. Thus, JOKER represents a step forward aiming at improving the capability of ad-hoc networks for transmitting heavy and demanding traffic such as multimedia content while reducing energy consumption in routing tasks. To the author's knowledge, there is not any prior work addressing the trade-off between the provision of quality for multimedia service and the reduction of node's energy consumption in opportunistic networks.

In the following, a wide overview about the trade-off between the quality achieved by multimedia services and the energy constraints imposed by mobile devices is presented. Then, it is discussed about the opportunistic routing paradigm, providing a brief review focusing on the most relevant proposals in this field during the last years. Thereafter, the novel opportunistic routing protocol JOKER is introduced and deeply described. Finally, the results section is divided in two well-differentiated subsections.

First it is provided a detailed evaluation focused on the performance of three prominent ad-hoc routing protocols from both QoE and energy-consumption perspectives. Then, a concise evaluation of JOKER supporting multimedia traffic in different-nature test-benches is presented.

5.2 Energy Consumption vs. QoE

The development of power-efficient techniques to diminish energy consumption in communication networks is a compelling need. In addition to global warming, economical interests, and other large-scale issues, there are other small-scale reasons for reducing energy-consumption levels when consuming networking services. These reasons are mostly related to the usability of end-devices, e.g., tablets or cell-phones. However, regarding multimedia services, it is a widespread belief that more energy is needed in order to achieve greater levels of quality. For that reason in this chapter it is presented a comparative study aiming at understanding the trade-off between the quality perceived when consuming a video-service and the energy consumed by the end-device. Concretely, the focus is on measuring the energy consumed by wireless cards, ignoring other sources of battery draining such as video (de) coding or other processing tasks. To this end, three different routing protocols, namely BATMAN (Better Approach To Mobile Adhoc Networking) [204], OLSR (Optimized Link State Routing) [207], and AODV (Ad hoc On-Demand Distance Vector) [209] are evaluated while managing an ad-hoc network with different number of video-streaming transmissions flowing through it. As discussed in previous chapters, BATMAN and OLSR are two of the most representative pro-active ad-hoc routing protocols, whereas AODV follows a reactive approach. This evaluation is performed under different environmental conditions and motion situations, evaluating for each protocol the attained QoE, in terms of MOS, for the video transmissions and the energy needed to accomplish this task.

5.2.1 Recent Advances in Energy-efficiency in Ad-hoc Networks

The popularity of video applications via streaming, conferencing, or sharing has put energy saving on the top of current demands for mobile and network devices. The Green Metadata standard [242], [243], currently in its final approval stage at ISO (International Organization for Standardization), is a good example of the highly importance of power reduction in video transmission, and how it is possible to achieve

it compromising QoE as less as possible from mobile devices perspective. In fact, exploring energy consumption versus QoE has produced its first results in several technologies and services such as Long Term Evolution, WiFi, or Wimax, among others [244]–[250].

Facing the QoE-and-energy aware video transmission from the physical layer, authors in [249] proposed to apply multicasting in an energy-efficient radio resource allocation scheme based on video quality as perceived by the end user. Their proposal was tested in an Orthogonal Frequency-Division Multiplexing (OFDM) scenario using a genetic algorithm as the basis of the allocation scheme. Results suggested that it is possible to steadily control energy requirements while maintaining high levels of perceptual quality as the number of subscribers increases in different multicasting scenarios. Nevertheless, the video quality method employed in this work, namely ITU-T Rec. J.144 [140], is outdated for video streaming in wireless environments and using the new ITU-T recommendations would be suggested to extend the validity of authors' proposal. Xu *et al.* [247] introduced two new performance metrics called QoE-aware Energy Efficiency (QEE) and QoE-aware Spectral Efficiency (QSE). Based on these parameters, which provide information about power consumption and bandwidth consumption from video quality's perspective, authors proposed a new spectrum and energy efficient mobile association and resource allocation scheme for wireless heterogeneous networks. In their work, video QoE is characterized in terms of Peak Signal to Noise Ratio (PSNR), and the solution proposed is formulate as an association and allocation optimization problem. After evaluating the system performance including PSNR distribution and resource utilization, they concluded that it greatly depends on the bandwidth and power decaying factors of cellular networks, leaving for further research the study of trade-off between QEE and QSE.

Signhal *et al.* [244], [245] advanced the state of the art by presenting a cross-layer optimization framework to enhance QoE and energy efficiency of heterogeneous wireless multimedia broadcast end devices. Their solution combines user composition-aware source coding rate (Scalable Video Coding) optimization, optimal time-slicing for layer coded transmission, and a cross-layer adaptive modulation and coding scheme. Simulation results in a single-cell video broadcast network using Digital Video Broadcast-hand-held (DVB-H) technology revealed that it is possible to group users based on their device capabilities and estimated channel conditions experienced by them

and to broadcast adaptive content to these groups, resulting in a better performance than current employed techniques.

Focusing on MANET, numerous works have studied and proposed energy-aware solutions for routing protocols (e.g., [251]–[253]). However, there are not related works to the author’s knowledge when it comes to evaluate both QoE and energy consumption. Only a few works can be found that aim to improve energy efficiency and QoS, i.e., quality from the classical network perspective, and most of them for on demand routing protocols [254]–[260]. As an example, Chen *et al.* [254] proposed a modification to the AODV protocol by introducing bandwidth estimation through the dissemination of bandwidth information in AODV hello messages. They obtained satisfactory results, although no node-movement was included in their evaluated scenarios. Other similar works are able to improve delay, packet delivery ratio, or jitter, but 1) the formal evaluation of quality of user experience is not included, 2) many of them are focused only on stationary nodes, and 3) most of them do not cover video streaming. With our work, we aim to fill this gap providing a first approach to QoE-energy design of routing protocols in MANET for video delivery.

5.3 Opportunistic Routing

In this section, the main foundations of the opportunistic routing paradigm are provided. As mentioned above, there are two well-differentiated phases in opportunistic algorithms. First, each node has to select a set of its one-hop neighbors as potential candidates to reach a given node. In this phase, usually known as candidate selection, the node also assigns priorities to its candidates according to some criteria. Once the candidate set is chosen, the transmitter includes the candidates IDs in the packet to be sent and transmits it to its neighbors. At the moment of receiving the packet, the involved candidates have to decide which of them will forward the packet in order to make the communication advance, avoiding duplicate transmissions. This phase is generally referred to as candidate coordination.

Regarding the candidate selection phase, several metrics to select and prioritize candidates have been proposed in the literature. Some of them are inherited from traditional ad-hoc routing metrics: hop count [261], geographical position [262], or expected transmissions count (ETX) [263]; others present more complexity and have been designed regarding the particularities of opportunistic scenarios; as an example,

Expected Distance Progress (EDP) [264] gathers the links delivery reliability and the nodes position to obtain a more consistent candidate selection metric. Another complex metric is that presented by Zhong *et al.* in [265]. In this work authors developed EAX (Expected Any-path Transmission), which is an evolution of ETX, considering the multiple paths that can be employed in opportunistic networks, too. Although this metric calculates quite efficient end-to-end routes, it is very costly to be computed in large networks due to its recursive nature.

In turn, a proper coordination among candidates is also crucial for the efficient forwarding of the packet in order to avoid unnecessary duplicated transmissions and to reduce the delay on the relaying operation. Specifically, in this stage the candidate with the highest priority that has received the packet should forward this message and the others have to discard it. One of the most intuitive candidate coordination methodology is the ACKnowledgment-based coordination [266]. Under this approach, when candidates receive a packet, they return an ACK message to the transmitter. When receiving the ACKs, the transmitter takes a decision taking into account different factors (highest priority, first ACK received, SNR, etc.) about the best candidate to relay the packet; once the best candidate is elected, the transmitter sends a “forwarding” control packet to it, authorizing the relaying of the packet. This methodology completely avoids duplicated forwarded packets at the expense of overloading the network with additional control messages, which also introduce more delay in the communication. Following this scheme, the RTS/CTS (Ready-To-Send/Clear-To-Send) coordination was proposed in [262]. Under this approach, when a node has a packet to send, it first broadcasts a RTS message with the candidates IDs in its header. According to the established priority, the candidates return a CTS message in a slotted manner (the first candidate answers after a SIFS period, the second candidate after a 2 x SIFS period, etc.). Finally, the transmitter starts the packet transmission after the reception of the first CTS. In addition, this CTS message allows the channel reservation for the transmitter. The RTS/CTS coordination presents the same advantages and drawbacks than the ACK-based coordination. The main difference is that in the latter the packet is sent first, and then begins the coordination phase, while in the RTS/CTS scheme this operation is made in reverse order. Besides, both methodologies (ACK-based and RTS/CTS-based) require that the off-the-shelf 802.11 network-card implementation has to be modified in order to enable the sending of ACK and CTS packets in different time slots

Aiming at avoiding the use of explicit control messages in candidate coordination, the timer-based and network-coding coordination schemes were also proposed. In the timer-based approach [267], when a group of candidates receives a packet, each of them sets a forwarding timer proportional to its priority ($t_{wait} \times (priority - 1)$), i.e., the first candidate immediately relays the packet, the second candidate waits during a t_{wait} period before forwarding the packet, and so on. During their waiting-period, candidates listen to the medium to hear if a higher-priority candidate has already forwarded the message. Therefore, a candidate relays a packet when its timer expires, but only if no other candidate with higher priority did it previously. This model is quite simple and does not inject extra control packets in the network. However, if a candidate cannot hear a higher priority candidate, there will be a duplicated packet. The idea of network coding [268] is avoiding the candidate coordination phase by conforming different linear combinations of several source packets into batches and send them as single transmissions to different forwarders. Thus, when the destination receives multiple batches, it will be able to decode and extract the single source packets. The main disadvantage of this scheme is the redundant information flooding the network because the same single packet can be coded inside of different batches. This could provoke an excessive use of the network resources and can impact on the overall throughput of the network.

5.3.1 Recent Advances in Opportunistic Routing

During the last years, several works proposing novel opportunistic protocols for ad-hoc networks have been presented. Although most of them are based on the opportunistic-routing foundations explained in the previous section, some works introduced interesting proposals worthy of being mentioned; for a more extensive review, please refer to [239].

Lin and Chen [269] presented two different spectrum-map-empowered opportunistic routing protocols for regular and large-scale cognitive-radio ad-hoc networks. In this work, two different state-of-art network architectures are gathered together: cognitive radio and opportunistic routing. Thus, after obtaining a spectrum map indicating the available spectrum with the geographic area, the proposed protocols make use of a packet-delivery strategy based on network coding. Therefore, as explained above, the candidate coordination phase is avoided. In turn, the candidates are chosen and ordered by using the end-to-end delay and the expected number of transmissions (ETX) metrics.

Additionally, for large-scale scenarios, geographical information is also considered for selecting the best forwarders towards the final destination. The proposed protocols showed improved performance (in terms of end-to-end delay) in comparison with other opportunistic protocols such as MORE (MAC-independent Opportunistic Routing & Encoding) [270] and a shortest-path greedy algorithm.

Following a similar approach, authors of [271] developed a routing protocol based on cognitive networking with opportunistic routing for wireless sensor networks. In this case, the candidate set selection criterion is the distance between each neighbor and the destination, enhanced with information about the network density. In turn, instead of using network coding, the RTS/CTS strategy is adopted for the candidate coordination. The proposed protocol was compared with a simple opportunistic spectrum access protocol and a geographic opportunistic routing protocol similar to GeRaF (GEographic Random Forwarding) [262], showing to be the most efficient in terms of throughput, delay and energy consumption in an indoor environment.

Another schemes using the network coding strategy are those presented in [272]–[274]. In the case of [272], a multicast protocol was proposed. Authors employed different strategies: (i) LP-based opportunistic routing structure, which defines the candidate set by taking into account the trade-off between opportunistic forwarding and the contention probability, (ii) opportunistic feeding, which permits different nodes to serve as cooperative sources for feeding other multicast receivers, (iii) fast batch moving, by which the original source can quickly start sending coded packets of the next batch, allowing the coexistence of different batches in the network, and (iv) inter-batch coding that permits improving the overall throughput by using RLNC (Random Linear Network Coding) coding operations instead of the traditional XOR coding scheme. Results showed how the proposed algorithm outperformed MORE [270] and Pacifier [275] in terms of energy efficiency, throughput, and response time. Likewise, works in [273], [274] proposed unicast and multicast algorithms, respectively, to improve the performance of real-time video communication in wireless networks. Both algorithms presented network-coding-based proposals aiming at improving the bandwidth utilization and throughput of the network. The provided results showed how the proposed protocols overcame the performance of other opportunistic routing algorithms. However, the test-benches employed in both works seems to be limited

considering that just 1 simultaneous video flow was streamed to evaluate static and pre-defined topologies composed by up to 7 nodes.

Gathering the concepts of network coding and content-based prioritization Seferoglu and Markopoulou proposed a more complete approach to opportunistic video coding for video-streaming taking into account the importance of video packets in network code selection [276]. By improving the application-level throughput, the proposed protocol permits reaching higher levels of quality (in terms of PSNR) than other routing protocols.

From a different perspective, Wu *et al.* studied the problem of selfish behavior in multi-rate opportunistic networks, i.e., the case in which one node could manipulate its input/output metrics to lead its neighbors to take routing decisions for benefiting itself [277]. This fact could be especially harmful in the presence of highly demanding services. In this work, authors introduced a new opportunistic routing protocol that ensures the faithfulness of each node in the network, maximizing the end-to-end throughput. By employing an incentive protocol, the presented scheme optimizes the network performance, maximizing each node's payoff, too. Besides, the overall network throughput is also improved with the presence of selfish nodes.

Work in [278] addressed the impact of link correlation in the candidate set selection. Thus, a link correlation aware metric was proposed aiming at enhancing the performance of the routing tasks by selecting the nodes with diverse low correlated links as forwarder candidates. The performance of the proposed protocol was evaluated in a wireless sensor network, showing a great efficiency against a correlation unaware protocol in terms of energy consumption, number of transmissions and delivery ratio.

Regarding energy efficiency, work in [279] presented an opportunistic routing protocol for minimizing the energy consumption of the nodes composing a fixed wireless sensor network. Focused on a multi-hop chain topology, the proposed algorithm takes into account two key metrics to select the candidate set: the distance of the sensor nodes to the sink and the residual energy of each node. The candidate-coordination phase is carried out following an ACK strategy: each candidate successfully receiving the data-packet replies an ACK to the sender after a given period determined by its priority. This operation is performed only if no other ACK from a higher-priority node has been overheard. After the ACK reply is completed, the data-

packet is forwarded. The proposed protocol was tested in both simulation and realistic test-benches, showing better performance, in terms of energy consumption, than GeRaF [262] and a Minimum Transmission Energy (MTE) protocol.

Finally, from a QoS perspective, work in [280] proposed to exploit the geographic opportunistic routing (GOR) for multi-constrained QoS provisioning in wireless sensor networks. Thus, reliability and end-to-end delay QoS constraints were considered. To this end, the proposed scheme makes use of both the packet reception ratio and the distance progress towards the final destination provided by each one-hop neighbor as metrics to configure the candidate set. The candidate coordination is performed by employing a similar ACK strategy than that followed in [279]. Results showed a great performance of the proposed protocol in terms of end-to-end delay, communication cost, and delivery ratio, which are greatly valued metrics in scenarios with strict QoS requirements.

Although the discussed works propose a wide range of performance improvements to different-nature services, there is still a lack regarding the trade-off between multimedia-services QoE and energy consumption in ad-hoc networks. Additionally, most of the proposed opportunistic routing protocols need to make deep modifications to the wireless-card protocol stack by modifying or replacing the MAC layer protocol (e.g., 802.11). Thus, the algorithm proposed in this chapter tries to cover these gaps by (i) improving the QoE of streaming services while reducing the energy consumed in routing tasks, and (ii) being a ready-to-use piece of Software compatible with other layers' protocols and without needing the modification of the off-the-shelf implementations of these protocols.

5.4 JOKER

JOKER (auto-adJustable Opportunistic acKnowledgment/ timEr-based Routing) is an opportunistic routing protocol that takes some of its features from the architecture of the BATMAN ad-hoc routing protocol. JOKER is a proactive algorithm, designed aiming at being as simple as possible in order to be suitable to work in a variety of devices with computational or energetic constraints. Besides, different protocol's configuration parameters are adjustable in order to permit JOKER to be highly adaptive to different network conditions. In the following, it is presented the structure and the operational insights of JOKER.

5.4.1 Main Characteristics

JOKER works between the link and the network layers of the OSI protocols stack. All the regular traffic in a node, as well as its own control messages, are (de)encapsulated, processed, and forwarded by JOKER. Thus, any protocol on higher layers (IPv4, IPv6, UDP, TCP, etc.) can run over JOKER without any modification. Additionally, as the node-addressing regarding routing tasks is based on MAC-addresses, there is no need for a network-layer addressing management system, e.g., IP, which simplifies the network setting-up operations. On the other hand, the link protocol also remains unmodified; so that, the original IEEE 802.11 protocol (or any other) can be used just by setting the network card to promiscuous mode, as JOKER is in charge of handling all the received packets (or frames).

All the traffic, except the routing control messages that have its own format as discussed below, is encapsulated with the JOKER header (please, see Fig. 5.2). The header includes the following fields:

- Packet type: informs about the nature of the encapsulated packet. As explained later, three types of packets have been defined: unicast, ACK, or Forwarding packet.
- TTL: Time To Live of the packet, defined as the number of hops that it is permitted to travel before being discarded. This value is initially set to the usual figure of 32.
- Packet-id: 4-byte code that uniquely identifies each packet. It is obtained by calculating the CRC-32 of the payload.
- Final destination address: MAC address of the packet's final receiver.
- Candidate x address: MAC address of each potential candidate to forward the packet. In total, there are $N_{candidates}$ candidates, which is a protocol configuration parameter.

This header is placed between those corresponding to the network and the link layers, adding $12 + 6 \cdot (N_{candidates} - 1)$ bytes of extra overhead. Note that the address of the candidate with the highest priority is not placed here but in the MAC header. This

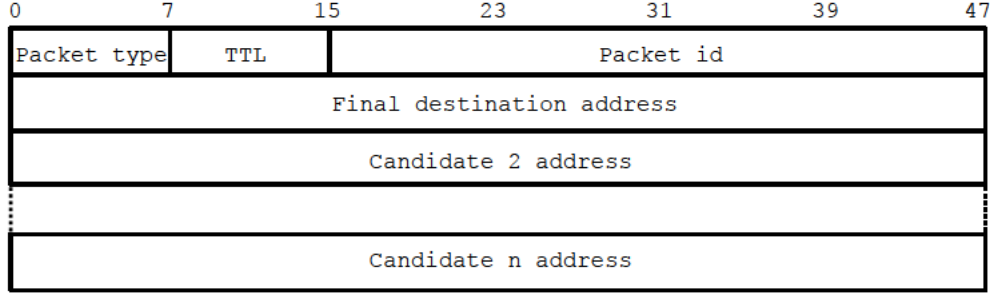


Figure 5.2. JOKER header.

fact will be deeply discussed in the next subsection. Also observe that, with the information included in the JOKER header, it is not necessary to pass the packet up to higher layers for routing or forwarding purposes. Furthermore, the inclusion of the packet's final destination address in the header allows the final receiver to accept and process lucky long transmissions. This situation, which was exemplified in previous sections (Fig. 5.1), reduces the transmission delay and avoids discarding valid packets in the receiving extreme of the communication.

5.4.2 Candidate Selection

As stated previously, JOKER takes inherited from BATMAN the way nodes interact with each-other regarding routing operations. Hence, it is kept the concept of transmitting a tiny packet, so called OriGinator Message (OGM) in BATMAN, as the unique control packet employed. By means of the reception and selective broadcasting of these messages, nodes are able to discover other nodes composing the network and establishing routes towards them. The original BATMAN metric made use of the number of OGM messages received from a given node via each link-local neighbor to estimate the quality of a route. In order to be able to find the best route to a particular end node, BATMAN counted the OGM messages received from each node in the network and logged which link-local neighbor relayed the greatest amount of messages. Currently, BATMAN's metric (transmission quality, TQ) has incorporated additional parameters regarding the asymmetry of the links and the number of hops in the routes towards the destination to obtain a more accurate assessment of the end-to-end path quality (see equation (5.1)).

$$TQ = TQ_{local} \cdot TQ_{recv} \cdot f_{asym} \cdot hop_penalty \quad (5.1)$$

TQ_{local} is the transmission quality locally calculated towards a direct neighbor to reach the final destination, TQ_{recv} is the transmission quality computed by the direct

neighbor towards the final destination, f_{asym} is an asymmetric links penalization, and $hop_penalty$ is a penalization for longer paths.

Previous works have demonstrated that BATMAN performs better than other ad-hoc routing protocols supporting real-time communications [220]. So that, the metric that JOKER employs to select and sort the candidate set is based on the BATMAN's metric but including a fine tuning. This adjustment regards the distance between nodes, penalizing the nearest neighbors against the most distant. The aim of this tuning is achieving a greater distance progress to the final destination in each hop, as the geographic routing protocols propose [262], [264]. Thus, the link quality (LQ) that a node assigns to a given direct-link to reach a distant node is calculated as showed in the equation (5.2),

$$LQ = TQ \cdot \frac{TQ_{max} - Distance_penalty}{TQ_{max}} \quad (5.2)$$

where TQ is the transmission quality calculated following the BATMAN algorithm (please, see equation (5.1)), TQ_{max} is the maximum value for the TQ (255 by default), and the $Distance_penalty$ factor is calculated based on the distance between the nodes, in other words, the progress that the neighbor provides towards the final destination. This value is estimated by assessing the fade margin (FM) of the packets received from the corresponding neighbor. The FM is defined as the difference between the power of the received signal and the sensitivity of the wireless card. Many routing protocols that make use of the link-strength as unique metric, select the best neighbors trying to maximize this value. These approaches tend to choose the nearest neighbors, which has the drawback of calculating routes formed by many short links. Adding more hops to the path causes additional transmissions, with its corresponding possible collisions, and adds more delay to the communication. For those reasons, the $Distance_penalty$ factor rewards the longest links against the shortest ones. According to different wireless-card manufacturers [281], [282], it has been defined the progress regions showed in Table 5.1. Three different regions have been established: high progress (low level of received power, $FM < 10$ dB), medium progress (medium level of received power, $10 \text{ dB} \leq FM \leq 20 \text{ dB}$), and low progress (high level of received power, $FM > 20 \text{ dB}$), with associated $Distance_penalty$ values of 1, 3, and 5, respectively. Observe that the tuning introduced by this factor is so mild that it does not have enough weight in the LQ calculation to modify very much the value of TQ . Thus, what it is achieved with this tuning is

TABLE 5.1. *Distance_penalty* mapping

Progress Towards the Final Destination	Fade Margin	<i>Distance_penalty</i>
HIGH	< 10 dB	1
MEDIUM	[10 - 20] dB	3
LOW	> 20 dB	5

selecting the most distant nodes among the best qualified candidates; therefore, JOKER multi- metric keeps the reliability of BATMAN's metric but also incorporates a beneficial factor rewarding those nodes that provide a greater progress to the end extreme of the communication without injecting additional control-packets in the network or making use of GPS information.

As mentioned previously, each node generates and broadcasts at a certain pace a simple packet to make the other nodes in the network know about its existence. Although the control messages' sending interval is usually fixed in almost all the ad-hoc routing protocols, as demonstrated in the previous chapter, the adaptation of this period to the actual network conditions is crucial for the well performance of the routing protocol; consequently, this also has a noticeable impact on the quality of the service flowing through the network. Thereby, JOKER incorporates a dynamic adjustment of the routing control-messages broadcasting time-interval. This period depends on the throughput that is momentarily crossing each node; hence, the nodes locally calculate its control message sending interval (CMSI) following the expression (5.3),

$$CMSI = 0.006 \cdot TP + 1.5 \quad (5.3)$$

where TP is the throughput crossing the node in Kbps. The validity of this dynamic tuning of the $CMSI$ is demonstrated in section 5.6.1.2, showing to be especially beneficial for networks loaded with heavy traffic such as those supporting multimedia transmissions.

Observe that the candidate selection is finally computed from a cross-layer perspective. The LQ value gathers network-layer metrics (end-to-end link reliability), link-layer metrics (link symmetry taken into account in the BATMAN's original metric), and physical-layer metrics (distance between nodes). Additionally, the dynamic $CMSI$ regards the application layer by adjusting the broadcasting of control traffic to the throughput generated by top-layer services. According to this multi-layer metric, JOKER selects the best $N_{candidates}$ as potential forwarders for a packet destined to a given

node. As mentioned above, $N_{candidates}$ is a protocol parameter and its impact on the performance of JOKER will be evaluated in following sections.

Finally, please note that the first candidate's address is not included in the JOKER header. Instead, it is passed down to the link layer as an input parameter, letting this layer to include this address in its header's most appropriate field. This is done aiming at both to avoid size-overspending in the JOKER header and to permit a wide compatibility with lower layer protocols. This also permits JOKER to work as a traditional ad-hoc routing algorithm, making use of just one candidate that will perform as the unique potential next hop in the path towards the final destination.

5.4.3 Candidate Coordination

In JOKER, two different candidate selection schemes can be employed, depending on the requirements of the service flowing through the network. For applications that do not have stringent delay-constraints and with some reliability needs, it has been designed the ACK-based candidate coordination methodology. Following this scheme, when the candidates receive the packet to forward, they extract its packet-id and return an ACK message with this identifier to the packet transmitter. ACK messages are sent as unicast packets without any candidates in the JOKER header in order to (i) reduce its size and (ii) avoid control packet storms. To reduce the delay in this process, there is not any priority rule for the order in which the candidates send their ACK messages, i.e., there is no distinction among candidates: once the packet is received, the ACK message is generated and sent. In turn, the packet transmitter answers with a Forwarding message packet (again, without candidates in the JOKER header) to the sender of the first received ACK. Therefore, this coordination scheme first makes use of the established candidate priorities computed by the candidate selection algorithm, and finally, it relies in the current state of the links between transmitter and candidates. In other words, it is selected the best candidate in the forwarding-operation moment among all the best potential candidates. Once the Forwarding message is received by the chosen candidate, it automatically forwards the data packet. As mentioned in the previous section, this methodology adds more control traffic and delay to the transmission; nevertheless, the number of duplicated packets is reduced to 0, and the election of the forwarder is made based on the freshest information about the status of the relaying nodes and the links between them and the transmitter.

In order to provide another candidate coordination methodology that met the requirements of heavy-traffic applications with low tolerance to delay, e.g., multimedia-content distribution, a timer-based coordination has been developed. By using this coordination scheme, when a candidate receives a packet, it waits a period of $t_{wait} \times (priority - 1)$ ms before forwarding the packet; thus, the highest priority candidate automatically forwards the packet once it is fully received and the rest of candidates listen to the medium before its forwarding timer is expired aiming at determining whether the packet has been forwarded by other candidates or not. The packet is only forwarded if no other previous packet-relaying is heard. As explained above the waiting timer can be tuned depending on the network and applications characteristics. This methodology does not inject any control packet into the network and reduces the extra delay of the ACK-based methodology. However, if a lower priority candidate does not hear the transmission of a higher priority candidate, duplicated packets can appear. The other drawback of this scheme is the extra delay added to the communication as the priority of the actual packet forwarded decreases.

5.5 Development Details & Test-Bench Description

JOKER has been coded to work in both simulation test-beds and real environments. Focusing in the former, a JOKER version in C++ language has been developed to be integrated in the InetManet Framework of the Omnet++ v.4.1.1 network simulator [10]; meanwhile, the JOKER version developed to work in real devices was coded in C, and has been tested in Ubuntu 12.04 (Linux Kernel 3.2.0) real computers. Due to the differences between the two test-beds considered, in the following, the characteristics of each test-bench are discussed separately.

5.5.1 Simulation Test-bed

The simulated scenario in Omnet++ was an IEEE 802.11g wireless ad-hoc network at 54 Mbps. The network was composed by 25 nodes with both static and dynamic mobility configurations. The nodes were randomly spread in a square area of 500 m x 500 m., thus having a node density of 100 nodes/km². Aiming at evaluating different-topology scenarios, 10 independent simulation instances with different seeds have been run for every protocol and scenario characterization. Hence, in order to generate the results, the average value for every measured parameter was taken, avoiding non-

representative singularities. The 95% confidence intervals for every measurement have been also calculated. Each simulation-run lasted 75 s.

The traffic pattern employed was generated by the InetManet Framework's UDPVideoServer2, which emulates the transmission of a video flow from a video server to an end-user, i.e., emulating a video streaming application. In this case, it has been employed a realistic 30 s video-trace available in [283], corresponding to a video encoded at 30 frames per second (fps) with the H.264/SVC single layer VBR (Variable Bit Rate) codec. As the considered devices are cell-phones, the video frame size was 176 x 144 px (QCIF format), the GoP (Group of Pictures) 16, and the number of B frames between I and P frames is 1 (G16B1). Given these video features, the encoding system generated VBR frames with an average size of 485.5 Bytes, transmitting them at 118 Kbps. With this video coding configuration, it was intended to simulate a handset-oriented video-streaming application. As the 802.11 MAC layer's MTU (Maximum Transmission Unit) was set to its default value (2304 Bytes), it was not necessary to fragment each video frame into different packets. The starting time for each multimedia communication was generated randomly according to a Poisson distribution in a time interval of (0, 10 s), after the convergence time for each routing protocol.

Aiming to obtain the most realistic results, the PHY layer has been simulated by means of the Nakagami- m propagation model, which considers the effect of fading in the wireless transmissions. The m factor has been set to 5. It has been also collected data when characterizing the physical layer with the Free Space propagation model. Thus, a comparison between the system behaviors in two different environmental hostility conditions has been performed.

As aforementioned, the network performance has been evaluated in both static and dynamic conditions. Focusing on the nodes movement, it imitates the motion of a pedestrian, following a Random Way Point mobility pattern with a speed characterized by a Gaussian distribution with a mean value of 1.34 m/s and a standard deviation of 0.26 m/s; the stop period was established to a uniform value between 2 s and 5 s [229].

Regarding the node's hardware configuration, it has been simulated the normal operation of a mobile handset. Thus, each node's wireless card has been configured following the real parameters of the widely used Broadcom's BCM4330 IEEE 802.11 a/b/g/n chip [284]. This chipset has been installed in different brand's cell-phones and

tablets, namely, iPhone 4s, iPad 3, Samsung Galaxy S 2, or Samsung Galaxy Tab 2, among others. Thereby, the transmission power was set to 20 mW (13 dBm), with sensitivity at 54 Mbps of -74 dBm and a SNIR (Signal to Interference and Noise Ratio) threshold of 4 dB. The employed energy-consumption model is based in that proposed by Feeney *et. al* [285]. This model registers the wireless card's state (transmit, receive, idle, or sleep), calculating the momentarily consumption according to the current state and the card's consumption associated to this state. In this case, the energy-consumption values presented in the BCM4330 chip's data-sheet [284] have been employed to characterize the simulated wireless cards (please, see Table 5.2). It was assumed that the battery's output voltage is 3.6 V. All the parameters and configurations of the link and network layer's protocols, namely, IEEE 802.11 MAC and IP, respectively, have been set to its values by default (Table 5.3). As mentioned above, JOKER does not need any modification in other layer's protocols for its properly functioning.

5.5.2 Experimental Test-bed

The evaluated scenario for the real implementation was similar to that showed in Fig. 5.1 (please, note that the link reliability values showed in that picture are only for the theoretical exemplification presented in previous sections and do not correspond to the actual delivery probabilities of the employed test-bench). This scenario was deployed by using the Emulab platform [9]. This network test-bench provides both wired and wireless nodes with high level of Software customization. The Emulab nodes' characteristics are showed in Table 5.4

In order to ensure that the TX node was not able to reach the RX node directly but by employing one of the relays (please, see Fig. 5.1), the wireless cards transmission power was set to 1 mW, since this is the minimum power value supported by the card [286]. The IEEE 802.11 channel 10, which was free of other transmissions during the experiment measurements, was employed, thus avoiding undesired interferences and collisions.

The test traffic used in the Emulab test-bench was generated by the Distributed-Internet Traffic Generator v2.8.1 [225], which allows sending video traffic by generating bursty traffic characterized by a given distribution and bit-rate. The video packets were fixed to 512 Bytes and the average throughput was set to 150 Kbps. Three independent and consecutive 60 s transmissions were established between TX node and

TABLE 5.2. Wireless card energy consumption

State	Current Consumption
Transmit	250 mA
Receive	60 mA
Idle	1.25 mA
Sleep	0.18 mA

TABLE 5.3. 802.11g and routing protocols parameters

Parameters	Bytes	Time
SIFS, DIFS, SLOT (μ s)	-	{10, 28, 9}
CWMIN (slots)	-	31
PLCP preamble (μ s)	-	4
{PLCP, MAC, SNAP} headers (μ s)	-, 28, 8	{16, 4.15, 1.18}
IP + UDP + RTP headers (μ s)	40	5.92
ACK (μ s)	14	2.07
BATMAN OGM interval (s)		1
BATMAN purge timeout (s)		200
OLSR {HELLO, TC, MID} intervals (s)		{2, 5, 5}
OLSR {Neighbor, Topology} hold times (s)		{6, 15}
AODV Active Route timeout (s)		3
AODV Hello interval (s)		1

TABLE 5.4. Emulab nodes characteristics

Feature	Value
Emulab PC type	pc2400w
Processor	Core 2 Duo E6600
Processor speed	2.4 GHz
RAM	2048 MB
Wireless Card	Netgear WAG311 802.11a/b/g (Atheros)
Operative System	Ubuntu 12.04. 64 bits
Linux Kernel version	3.2.0
BATMAN version	Batman-adv 2014.2.0
OLSR version	olsrd 0.6.6.2

RX node. 5 independent tests were performed in order to avoid non-representative results, calculating the 95% confidence intervals for every measurement, too. These values are not shown due to their low significance. Finally, aiming at emulating disconnection issues, the relay's network cards (node 1, node 2, and node 3)

have been randomly switched off during 5 second-periods during the transmission between the TX node and RX node. The beginning time from each disconnection event was scheduled following a uniform distribution from 30 s to 60 s.

Common to both test-beds, it has been employed QoE and QoS metrics to evaluate the performance of the routing protocols. Focusing in the former, the video QoE has been measured in terms of MOS (Mean Opinion Score) following the model presented in [145] and the corrections to this model introduced in [146]. This model allows obtaining accurate video-MOS estimations by considering the Packet Loss Rate (PLR), the coding scheme (k), the resolution (a) and the video-coding bit-rate (Br) as shown in expression (5.4).

$$V_q = 1 + 4k \left(1 - \frac{1}{1 + \left(\frac{a \cdot Br}{v_1} \right)^{v_2}} \right) e^{-\frac{PLR}{v_3}} \quad (5.4)$$

Besides, v_1 , v_2 , and v_3 , are experimental factors that depend on the coding scheme, the resolution applied, and the motion characteristics of the video. In this case, as described above, the widely extended H.264 codec has been used for the video-streaming transmission in QCIF format. It was assumed a video-content with low movement, as it is usual in video-telephony calls. Thus, the values assigned to these factors were: $k = 1.12$, $a = 10.8$, $v_1 = 0.366$, $v_2 = 1.32$, and $v_3 = 3.5$ respectively, hence consistent with the figures recommended by the model's authors [145], [146].

From the QoS perspective different metrics have been considered to accurately assess the network performance, namely, number of hops between video server and video client, Packet Delivery Ratio (PDR), throughput, and Packet Loss Ratio (PLR). The energy consumed by the wireless cards in routing tasks has been taken into consideration, too.

5.6 Results

In the following, the outcomes extracted from the described test-benches are presented. As discussed above, first we present a comparison study conducted in the simulation environment aiming at evaluating the energy consumed by three prominent ad-hoc routing protocols (BATMAN, OLSR, and AODV) supporting a video-streaming service under different environmental conditions. Besides, the quality attained by the multimedia service in terms of QoE is also considered. Thus, the trade-off between the

quality perceived when consuming the video-service and the energy consumed by the end-device are analyzed.

After the evaluation of these standardized protocols, it is proposed an enhancement to the BATMAN algorithm, not just aiming at improving the quality of the video streaming service, but also trying to reduce the energy consumption of mobile devices. In the previous chapter, it was demonstrated the need of properly tuning the time-interval between BATMAN's OriGinator Message (OGM) broadcasts (hereinafter OGM interval) in order to adapt this routing protocol to current network conditions. Thus, a characterization of the dynamic update of the OGM interval is provided here, presenting a mathematical formulation that permits to fit this control-packets sending-interval to the network load in real-time. By means of this enhancement, the BATMAN routing protocol turns to be more flexible and suitable to better support different types of traffic. All the results presented here have been extracted from the simulation test-bed, too.

Then, a performance evaluation of JOKER in terms of QoE, QoS, and energy consumption perspectives is carried out. This study includes results for several JOKER parameter characterizations in order to analyze their impact on the energy consumed by the wireless cards as well as the quality of a highly demanding multimedia service such as video streaming. These outcomes, extracted from both simulated and realistic test-benches, are compared with those attained by BATMAN.

5.6.1 QoE - Energy Consumption Trade-off Evaluation

In this section, it is provided a concise performance evaluation focused on the performance of the three routing protocols under study (BATMAN, OLSR, and AODV) from both QoE and energy-consumption perspectives. Results include stationary and non-stationary scenarios. Then, it is introduced a modification of the BATMAN protocol and its corresponding performance.

5.6.1.1 QoE vs. Energy Consumption Study of BATMAN, OLSR, and AODV

Fig. 5.3 depicts the average QoE (in terms of MOS) obtained for different number of simultaneous video streams flowing through the afore-described static network managed by BATMAN, OLSR, or AODV. The average power consumed by each node during each entire simulation run (75 s) is also shown. Recall that these power-

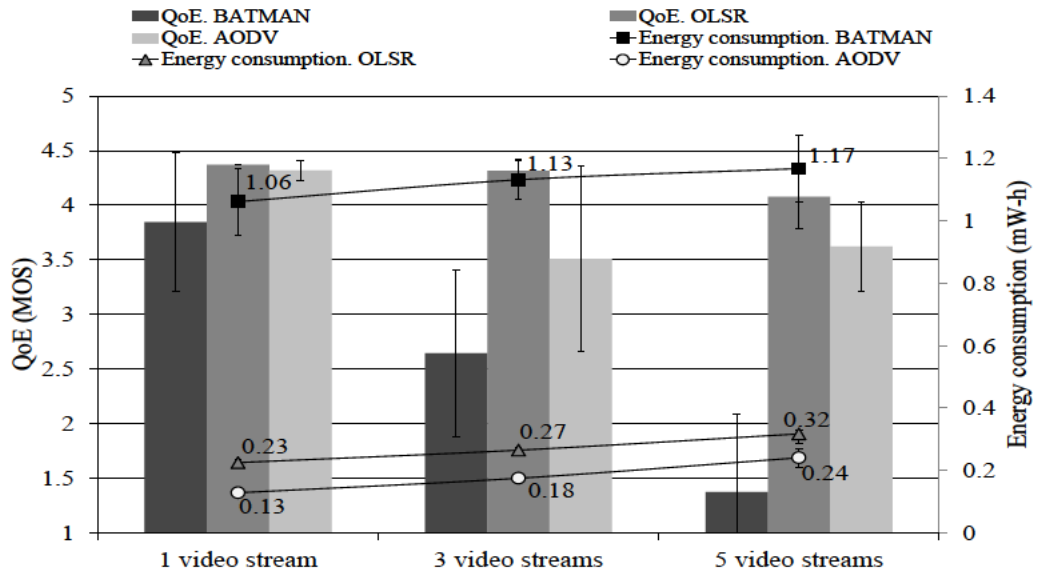
consumption values correspond only to wireless card consumption, ignoring other sources of battery draining. The results attained for two different levels of transmission-medium hostility are presented. Thus, Fig. 5.3 (a) corresponds to the evaluation under ideal Free Space conditions, whereas Fig. 5.3 (b) presents the results when characterizing the physical layer with the Nakagami- m propagation model (hostile conditions). Focusing on quality outcomes under different wireless channel conditions, the BATMAN protocol is the one with the most stable behavior. Observe that, although in the Free Space scenario (Fig. 5.3 (a)) OLSR and AODV present the best results in terms of QoE (MOS always above 3.5), in the more realistic Nakagami- m scenario (Fig. 5.3 (b)), the performance of these protocols dramatically drops, leading to QoE (MOS) values below 1.5. These results can be explained from the information-dissemination strategy adopted by each protocol: Both OLSR and AODV make use of few control packets containing a big amount of information inside them; therefore, packet loss affecting these routing-data packets severely harms the respective path-finding algorithm. On the other hand, BATMAN uses a less controlled flooding algorithm [204]: based on the periodical (re) broadcasting of tiny OGM packets, the routing information gets finally distributed among all network nodes.

These strategies are also reflected in the energy consumed by each algorithm. Observe that, due to the continuous broadcasting and relaying of control packets, the BATMAN protocol is the one with the highest power consumption. In turn, OLSR has a more controlled routing-packets dissemination due to its characteristic Multi-Point Relaying (MPR) strategy [207]; thus, it presents considerably lower energy consumption than BATMAN. Finally, AODV is the protocol that employs less control packets because of its reactive nature. This is clearly noticeable in its power consumption, being the lowest one among the three protocols.

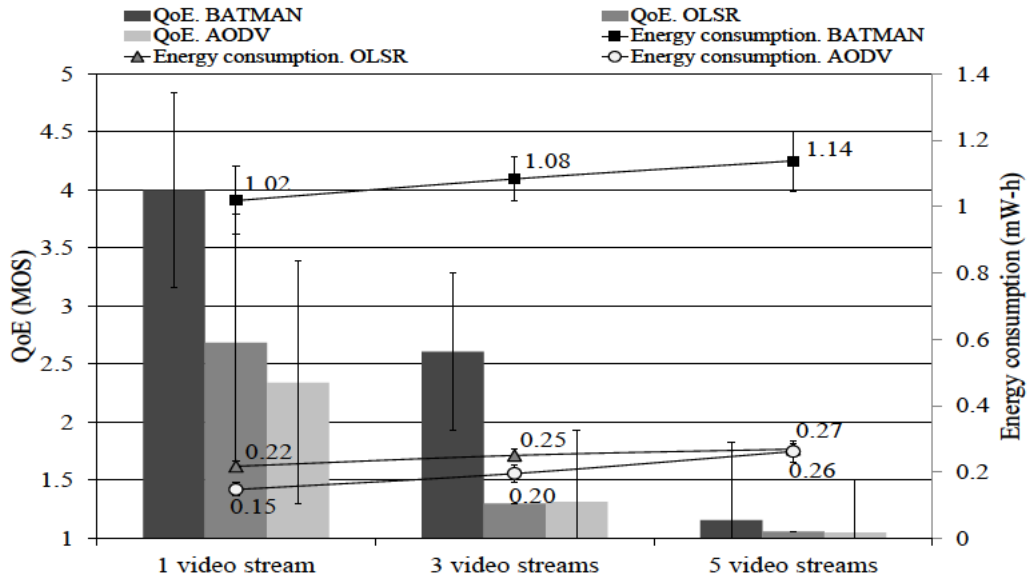
It is interesting to analyze the different behavior of each protocol's power-consumption pattern, depending on the hostility of the wireless transmission channel. While the proactive protocols (BATMAN and OLSR) consume more energy in the Free Space scenario, the reactive protocol (AODV) drains more energy in the hostile situation (please, compare Fig. 5.3 (a) and Fig. 5.3 (b)). This performance is explained as follows: BATMAN and OLSR make use of an information-dissemination strategy based on the periodical (re) broadcasting of control packets. Thus, as in the Free Space environment few broadcast-packets are lost, a greater number of re-broadcastings occur,

increasing the energy consumption compared to the lossy Nakagami- m scenario. In turn, the AODV's route discovery mechanism makes nodes massively re-broadcast received route request (RREQ) packets until a route to the destination is established [287]. Thereby, in low-loss scenarios, i.e., Free Space, the routes are established quicker than in the hostile ones, so less control-packet broadcastings are necessary, hence reducing energy consumption.

In the lights of these results, it can be concluded that in environments with low level of packet loss, i.e., Free Space situations, there is not a clear trade-off between the



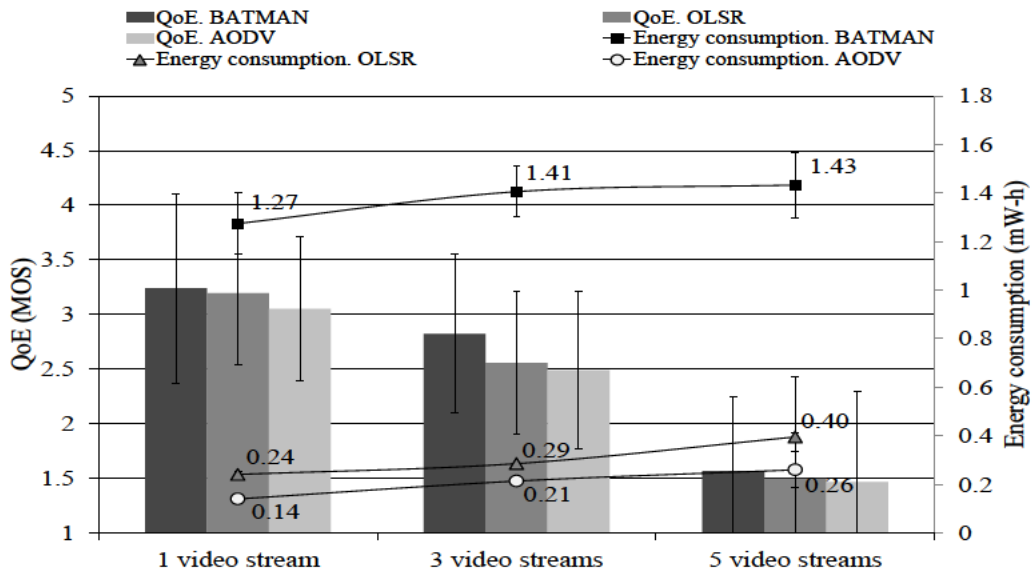
(a) Free Space conditions



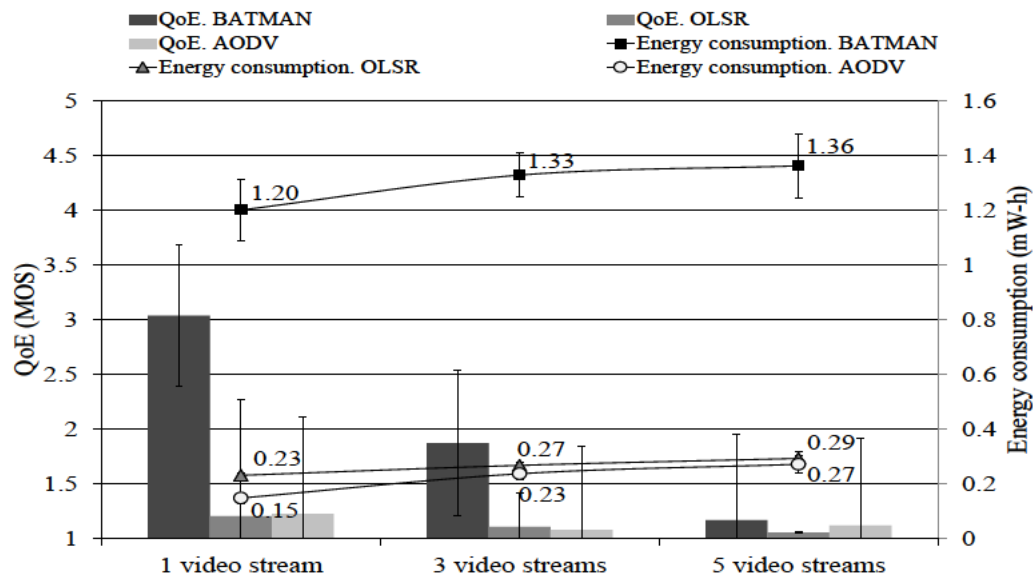
(b) Nakagami- m conditions

Figure 5.3. QoE vs. Energy consumption in static scenario for a variable number of simultaneous video streams.

quality attained by the multimedia communication and the needed energy for routing tasks. This statement is evidenced observing the QoE results attained by AODV, which are quite acceptable in terms of MOS compared with those obtained by the other routing protocols, while consuming low amount of energy. In lossy environments, our results suggest that the information-dissemination strategy of OLSR and AODV as a control-packet saving-strategy is clearly inefficient in this type of hostile scenarios, so other more energy-demanding routing approaches should be taken aiming at achieving higher levels of quality.



(a) Free Space conditions



(b) Nakagami- m conditions

Figure 5.4. QoE vs. Energy consumption in dynamic scenario for a variable number of simultaneous video streams.

Regarding the protocols' capability to manage continuous route re-calculations, Fig. 5.4 shows the same evaluated parameters as before but under dynamic conditions. As described previously, all nodes follow a motion-pattern imitating the movement of a pedestrian handling a cell-phone [229]. Firstly, please observe the overall quality decrease affecting the three protocols (Fig. 5.4) compared to the static scenario (Fig. 5.3). Again, BATMAN is the protocol showing the greatest strength against hostile conditions. Both OLSR and AODV show much more degraded performance, particularly noticeable in the Free Space environment (Fig. 5.4 (a)). In the Nakagami- m scenario (Fig. 5.4 (b)), node motion additionally contributes to the already mentioned poor performance of these protocols under lossy conditions.

Focusing on energy consumption, there is a logical increase in the power consumed by the three protocols under study. Due to the continuous changes in the routes between transmitters and receivers, the routing tasks become more intense and hence the energy draining. The protocol's energy consumption behavior regarding the transmission-channel hostility is similar than that attained in the static scenario: proactive protocols, i.e., BATMAN and OLSR, drain more energy under Free Space conditions, whereas AODV (reactive protocol) consumes more in the lossy scenario.

5.6.1.2 Improving BATMAN performance with dynamic OGM intervals: Dyn-BATMAN

By default, BATMAN OGM interval is a fixed configurable time lapse, with a recommended value of 1 s following the standard guidelines [204]. However, it has been discussed in the previous chapter about the need of properly tuning the BATMAN OGM interval in order to adapt the rate of control-packet broadcasts to current network conditions. This idea has been already applied to other routing protocols under consideration in this study [288], [289]. Nevertheless, to the author's knowledge, it has not been presented yet any dynamical adaption for the BATMAN routing protocol. In the previous chapter, it was demonstrated that an excessive number of OGM packets in the network noticeably harms the quality of the service flowing through it, especially in low-loss environments, e.g., Free Space. Thus, it was revealed that as the throughput in the network increased, it was necessary enlarging the time-interval between OGM broadcasts to maintain the desired quality level on the multimedia service crossing the network.

Consequently, it is proposed here to update the BATMAN Control-Message Sending Interval (CMSI) using equation (5.5) in order to control the number of OGM packets in the network. This new expression to update CMSI only employs the local information available in each node.

$$CMSI = a \cdot Throughput + b \quad (5.5)$$

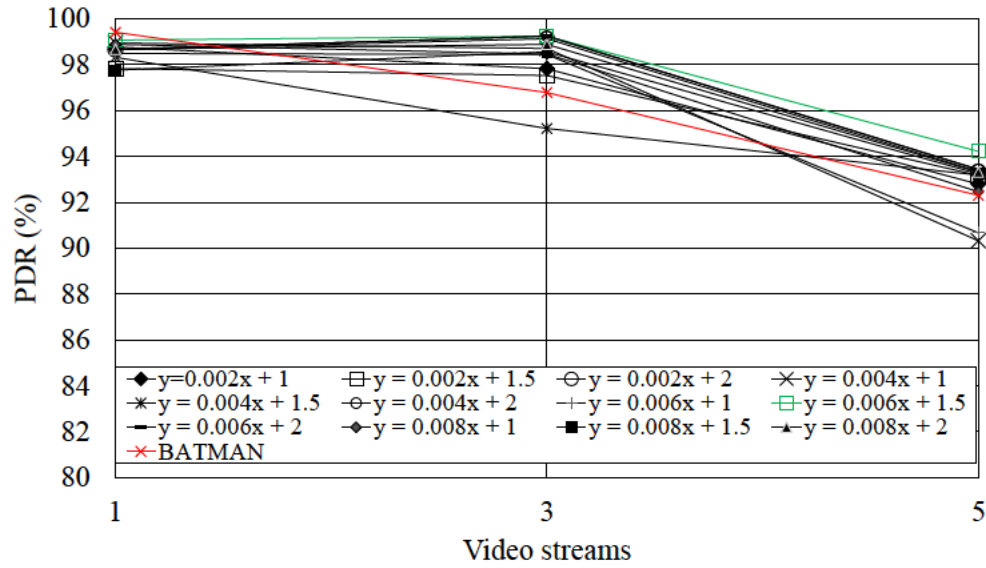
In equation (5.5), a and b are experimental shaping factors and *Throughput* is the traffic crossing the node (in Kbps) since the last CMSI update. The CMSI is given in seconds. In order to obtain the values for the shaping factors a single-target experimental loop optimization strategy has been followed. To this end, extensive computer simulation has been carried out. Packet Delivery Ratio (PDR) has been selected as target function to maximize. After several preliminary experimental analyses, the considered range of values for factor a was [0.001, 0.008]; in turn, the selected values for the factor b were 1, 1.5, and 2. Fig. 5.5 and Fig. 5.6 shows the PDR attained in the scenarios under consideration when applying the dynamic CMSI with the aforementioned ranges of shaping values. For the sake of simplicity, we only include here the results obtained with even values for factor a . For comparison purposes, the PDR values obtained by the original BATMAN have been also included (red line in Fig. 5.5 and Fig. 5.6). Thus, observe that in all evaluated scenarios, a proper tuning of the BATMAN CMSI leads to an increase in the average PDR in comparison with the original algorithm. Although the number of a and b factors combinations is high, there is one that stands out from the rest:

$$CMSI = 0.006 \cdot Throughput + 1.5 \quad (5.6)$$

This function has been marked in green in Fig. 5.5 and Fig. 5.6. By adjusting the dynamic CMSI's values to those showed in expression (5.6), the minimum time-interval between OGM's broadcasts is 1.5 s. This occurs when the throughput crossing the node is 0. Considering that the BATMAN's default OGM interval is 1 s, this enlargement of the minimum CMSI means a 33% reduction in the number of OGM packets in the network. Therefore, when applying the dynamic update of the CMSI, a controlled reduction of routing packets is achieved, resulting in an improvement on the performance of the service flowing through the network.

After selecting the most appropriate values for the CMSI, the performance of the dynamic BATMAN (Dyn-BATMAN) routing protocol needs to be compared with that

of the original one. Thereby, Fig. 5.7 presents a comparison between the QoE (in terms of MOS) and the energy-consumption attained by the Dyn-BATMAN and the original BATMAN in static scenarios. Results for AODV have been also included as it previously showed the best trade-off between QoE and energy consumption in some of the considered scenarios (please, see Fig. 5.3 and Fig. 5.4). Observe that in all simulated scenarios, the QoE (MOS) achieved by using Dyn-BATMAN is greater than that obtained by managing the network with the original BATMAN.



(a) Free Space conditions

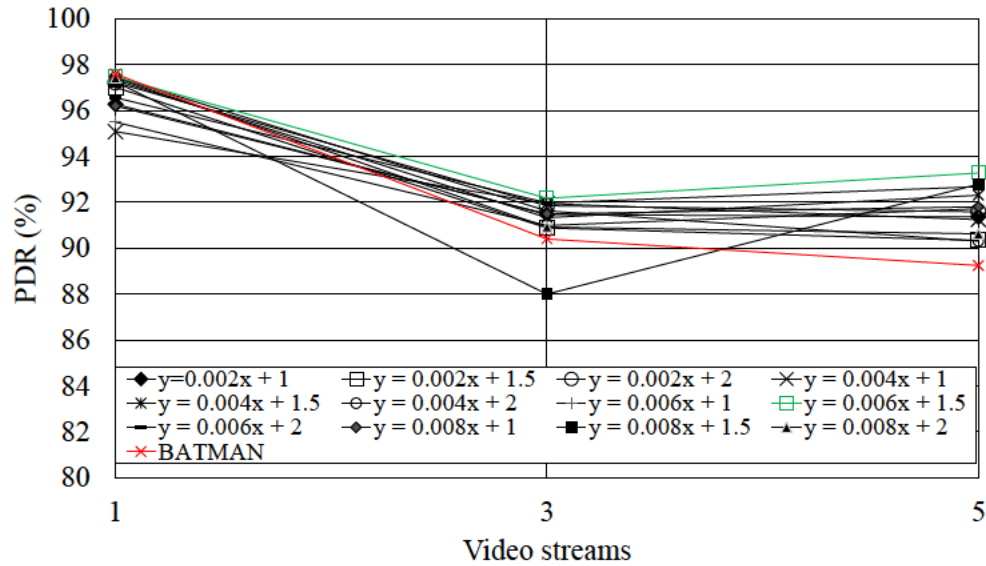
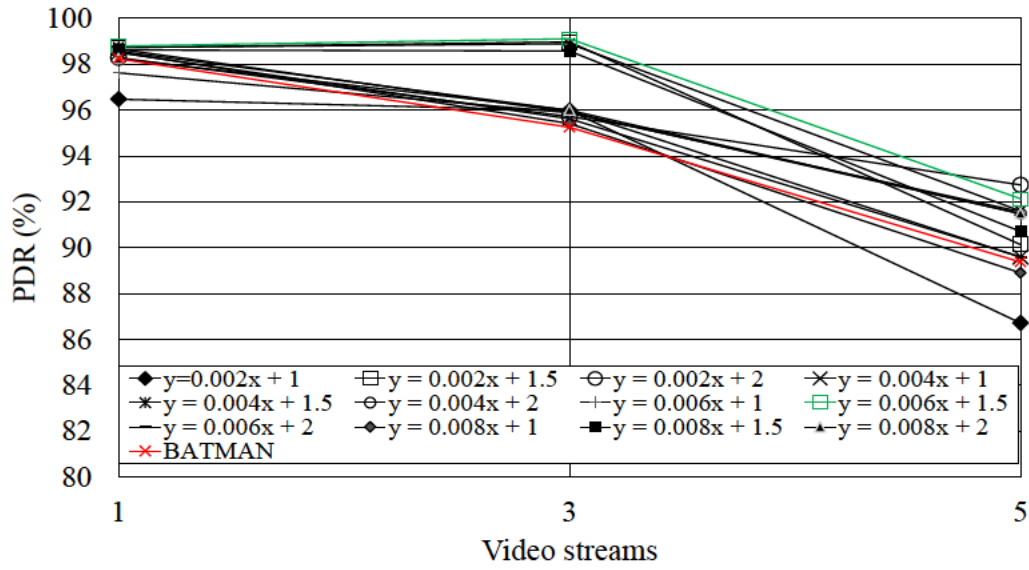

 (b) Nakagami- m conditions

Figure 5.5. Packet Delivery Ratio attained by the different adjustments applied to BATMAN CMSI in static conditions for a variable number of simultaneous video streams. Free Space (a) and Nakagami- m (b) propagations models.



(a) Free Space conditions

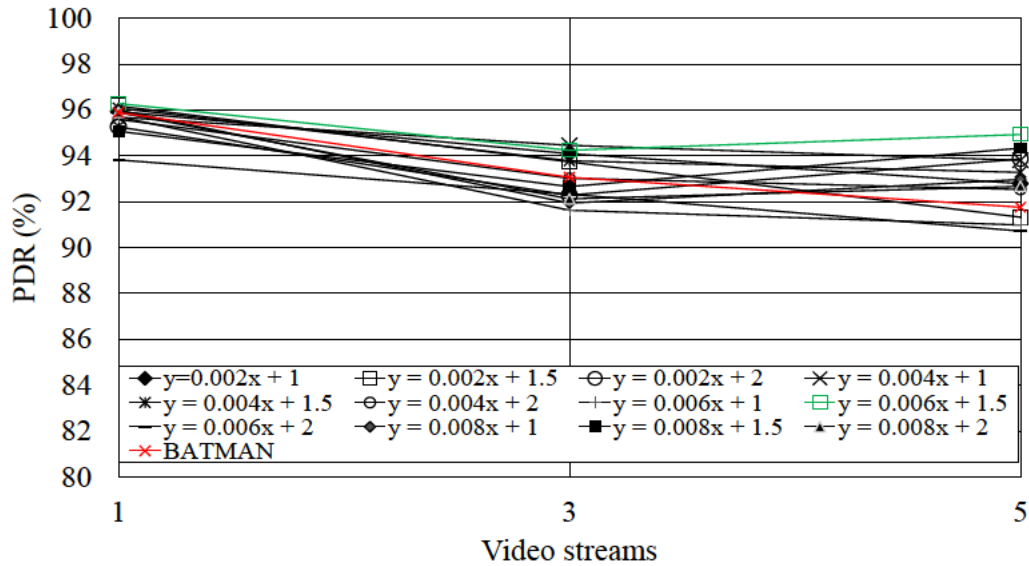
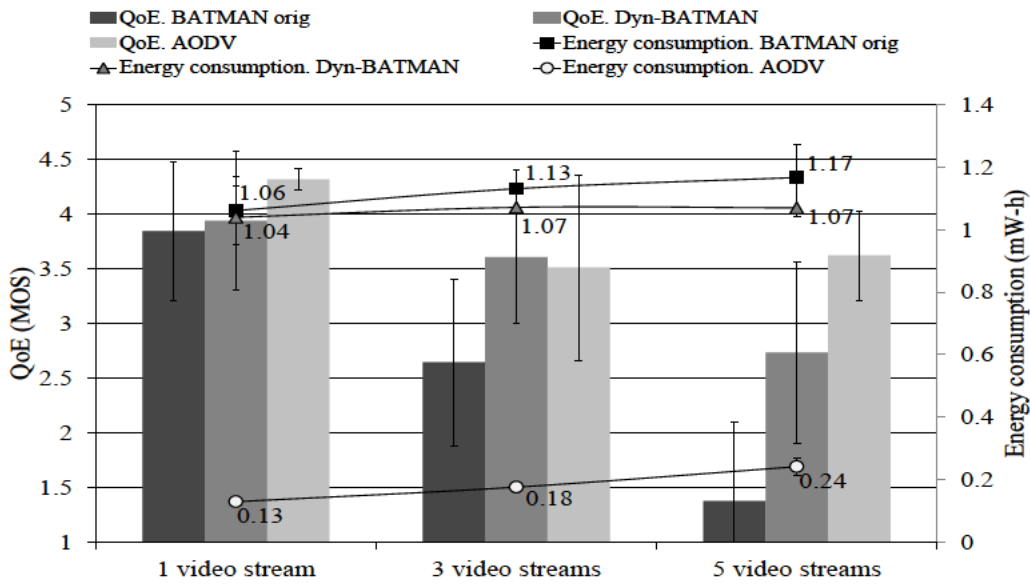
(b) Nakagami- m conditions

Figure 5.6. Packet Delivery Ratio attained by the different adjustments applied to BATMAN CMSI in dynamic conditions for a variable number of simultaneous video streams. Free Space (a) and Nakagami- m (b) propagations models.

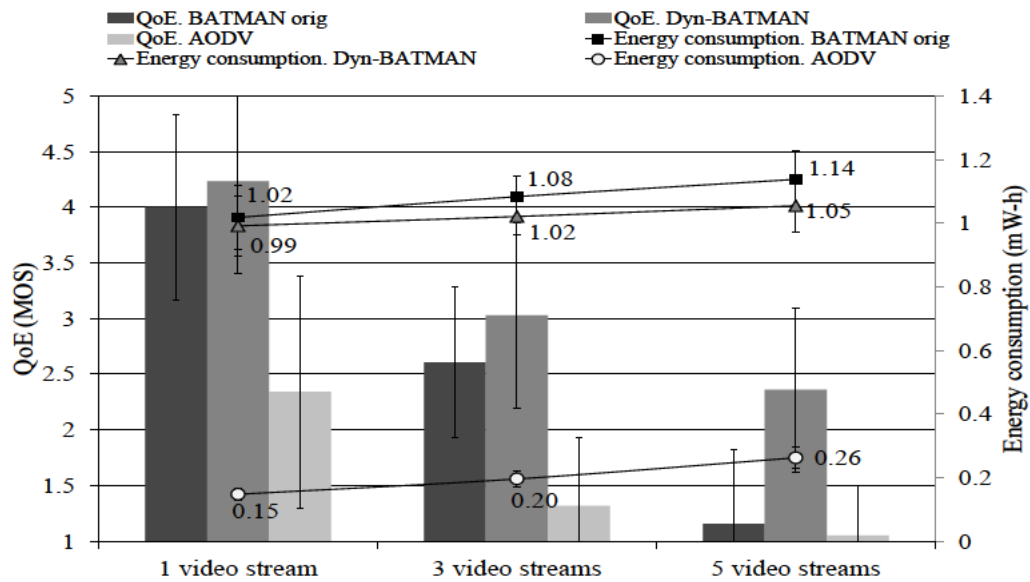
In the scenario characterized by the Free Space propagation model (Fig. 5.7 (a)), the superior performance of Dyn-BATMAN with 3 simultaneous video streams in the system is noticeable. Even so, under these ideal conditions, AODV maintains its superiority (in terms of MOS) in the most over-loaded scenario (5 simultaneous video streams). However, it is in the most realistic Nakagami- m scenario (Fig. 5.7 (b)) where Dyn-BATMAN clearly presents the best performance compared with the other protocols; observe how Dyn-BATMAN shows its robustness to the adverse propagation

conditions since it is the only protocol that reaches an acceptable QoE (MOS) value of 3 with 3 simultaneous video transmissions in the system. Under the heaviest traffic load, Dyn-BATMAN is the one that attains the highest QoE (MOS) level, although not sufficient to be considered as acceptable (QoE (MOS) below 3).

Regarding battery consumption, observe that in all simulated scenarios the power drained by Dyn-BATMAN is lower than that consumed by the original BATMAN, far from the lowest values achieved by AODV though. AODV is still the most battery-

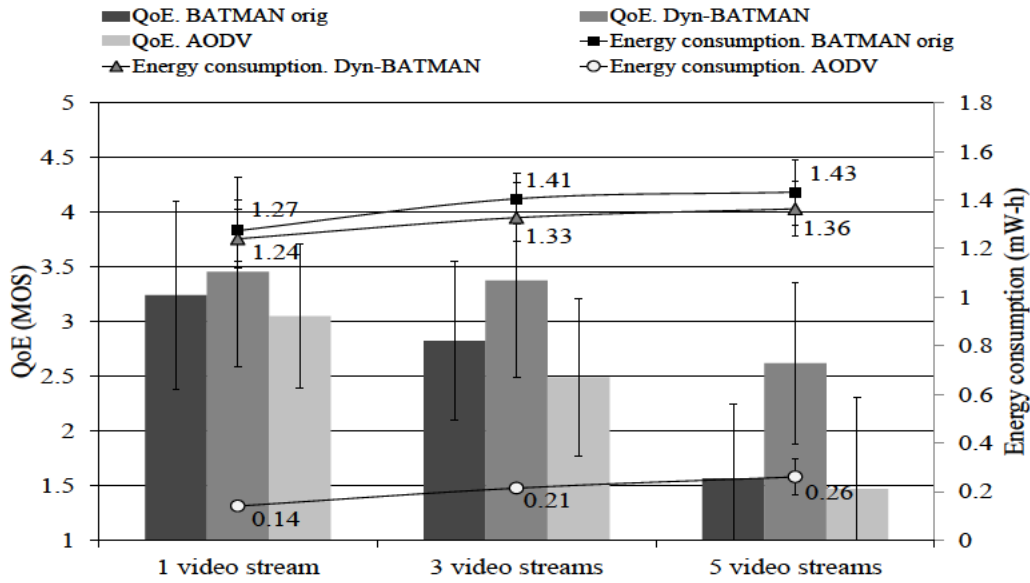


(a) Free Space conditions



(b) Nakagami- m conditions

Figure 5.7. Comparison of QoE and energy consumption among BATMAN orig, Dyn-BATMAN, and AODV in a static scenario with a variable number of simultaneous video streams.



(a) Free Space conditions

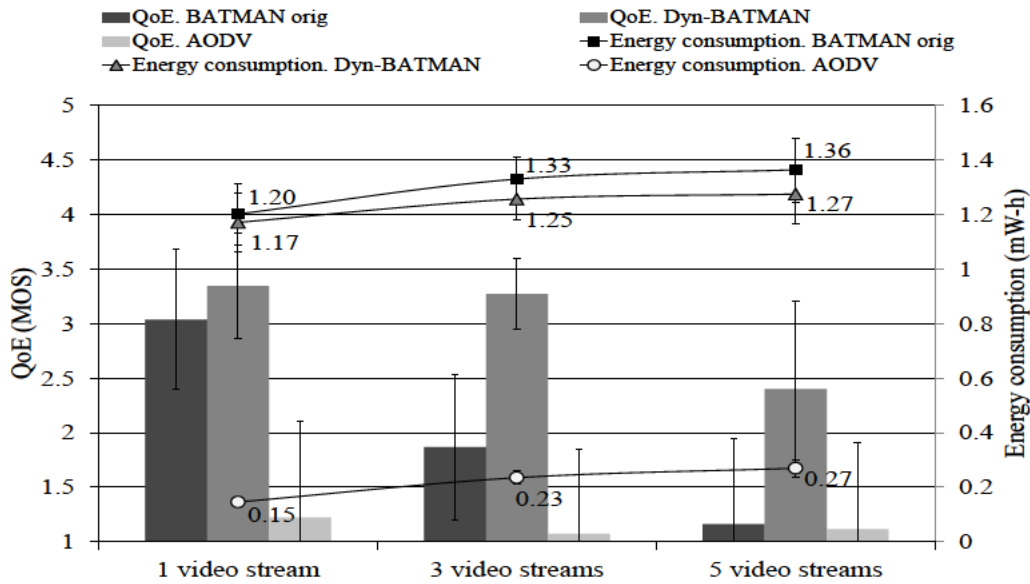

 (b) Nakagami- m conditions

Figure 5.8. Comparison of QoE and energy consumption among BATMAN orig, Dyn-BATMAN, and AODV in a dynamic scenario with a variable number of simultaneous video streams.

friendly protocol due to its reactive nature. The Dyn-BATMAN's power-consumption reduction is due to the decrease of OGM broadcasts by the proper enlargement of the CMSI. By applying this enhancement, wireless cards remain in an active state (transmitting or receiving) less time, entering in low-energy-consumption mode.

This performance evaluation described for the static scenario (Fig. 5.7) is also valid for the dynamic one (Fig. 5.8). In fact, Dyn-BATMAN performs even better under node-motion conditions compared with the other 2 protocols. Note that the highest

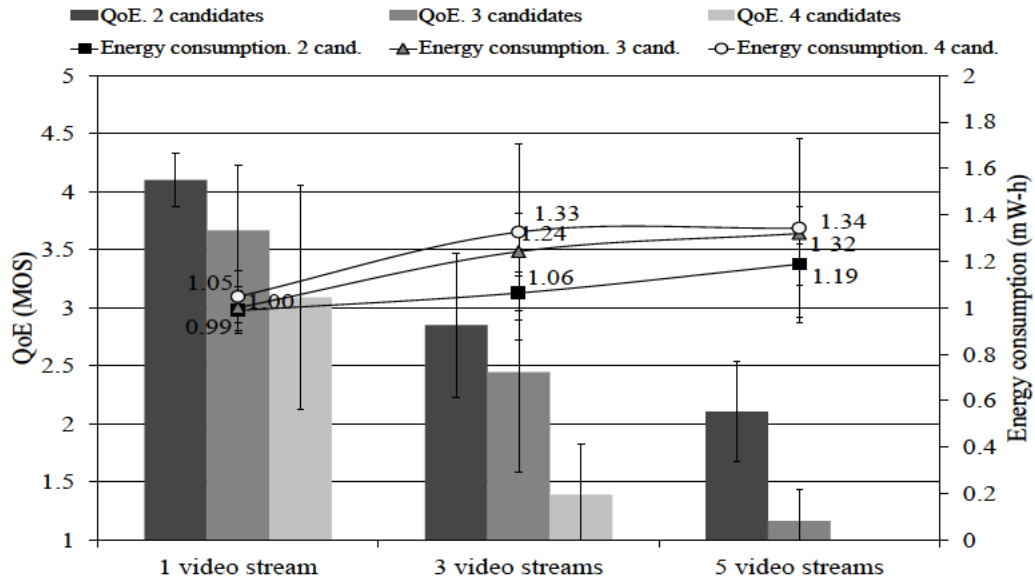
levels of QoE (MOS) are always achieved by managing the network with Dyn-BATMAN, no matter the number of simultaneous transmissions in the network or the wireless-channel conditions (Free Space, Fig. 5.8 (a), or Nakagami- m , Fig. 5.8 (b)). Although the QoE (MOS) values present a little decrease in comparison with the static scenario, the robustness showed against the difficulties posed by nodes motion is greater than that shown by the other protocols. Besides, the energy consumption remains under the levels achieved by the original BATMAN. Although the consumption of AODV continues in much lower figures than that obtained by Dyn-BATMAN, the quality level achieved by managing the network with AODV is now much lower than that attained by Dyn-BATMAN, too. Therefore, it can be concluded that the enhancement of Dyn-BATMAN in comparison with the original version of the protocol is twofold: (i) it permits an increase in the quality levels of the service crossing the network, together with (ii) a reduction of the energy consumed by the wireless cards. Thus, it has been demonstrated that a more efficient strategy of the ad-hoc routing protocols is possible, as it is feasible to obtain more battery-friendly routing algorithms without harming, even increasing, the quality level of the services flowing through the network.

5.6.2 JOKER Evaluation

In this section, the performance of JOKER supporting video-streaming traffic is explored. As aforementioned, many of the proposed protocol's parameters are tunable; for that reason, in the following different subsections showing the response of JOKER when varying these values are presented. All these outcomes are compared with those attained by BATMAN. The presented results for this complete evaluation have been extracted from the simulation test-bed due to its easiness for tuning the environmental and network conditions. At the end of this section, additional outcomes extracted from the realistic test-bench are also exhibited aiming at showing the functionality of the JOKER implementation for real machines.

5.6.2.1 Protocol Adjustment

Fig. 5.9 and Fig. 5.10 shows the average QoE (MOS) and the average wireless card's energy consumption for a variable number of simultaneous video streams transmitted through the network. In this case, the network was managed by JOKER, employing the ACK coordination (in the following JOKER-ACK). The results have been extracted from different scenario configurations, namely, static and dynamic, characterizing the



(a) Free Space conditions

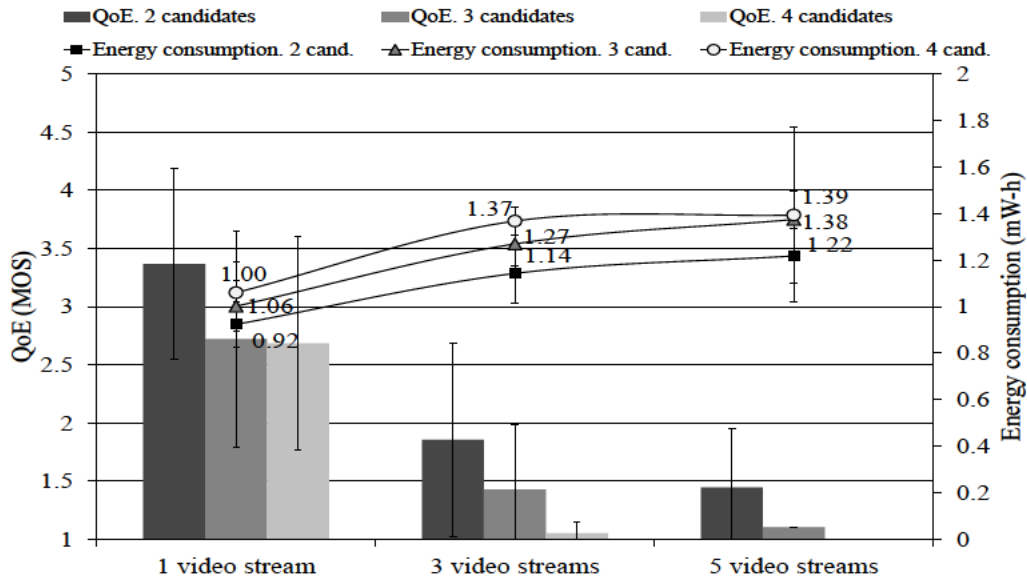
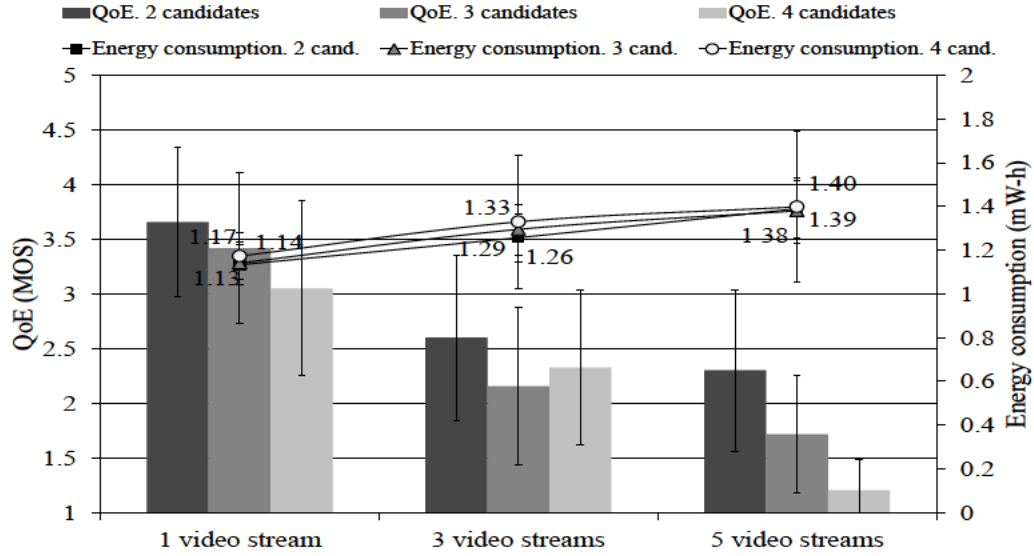

 (b) Nakagami- m conditions

Figure 5.9. Comparison of QoE (MOS) and energy consumption attained by JOKER-ACK for a variable number of simultaneous video streams and different number of opportunistic candidates. Results attained in static conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

wireless transmission channel with the Free Space and Nakagami- m ($m = 5$) propagation models. It is clear that the best results in terms of both metrics under consideration (QoE and power consumption) are obtained when considering just 2 opportunistic candidates, no matter the scenario conditions. This fact has sense considering that the ACK coordination injects more packets in the network. Therefore, when more candidates are involved in the routing activities, more ACK packets are



(a) Free Space conditions

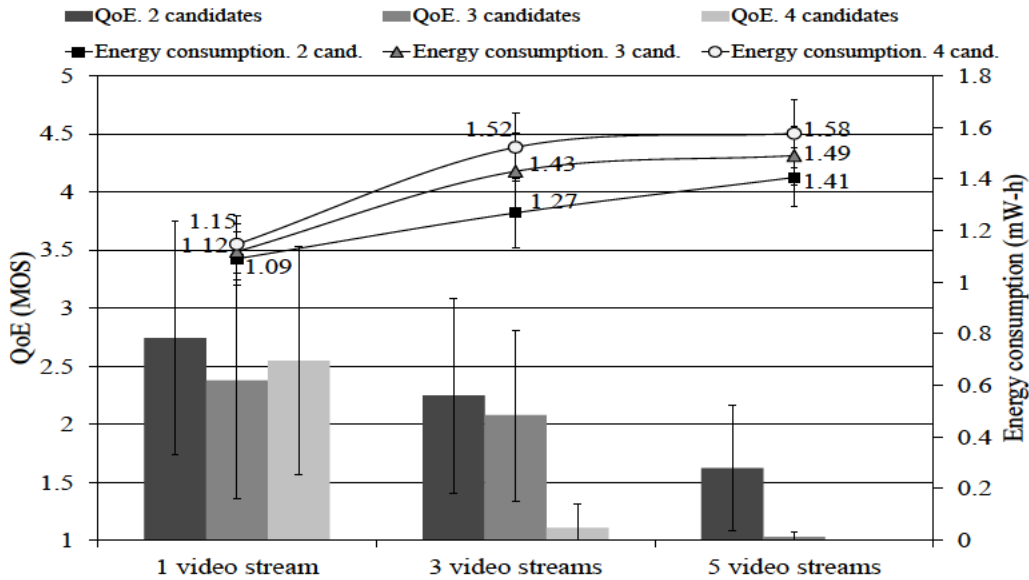

 (b) Nakagami- m conditions

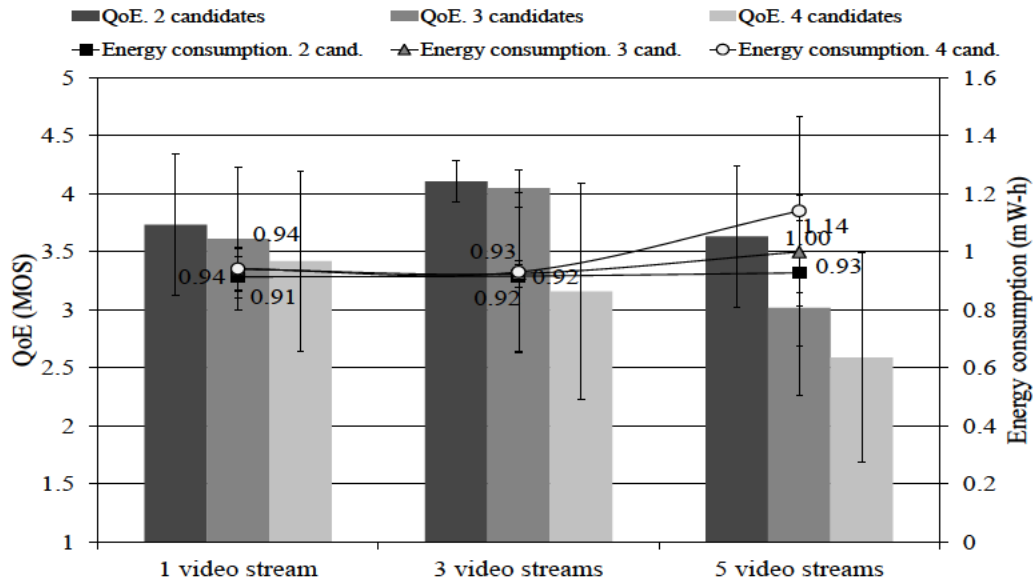
Figure 5.10. Comparison of QoE (MOS) and energy consumption attained by JOKER-ACK for a variable number of simultaneous video streams and different number of opportunistic candidates. Results attained in dynamic conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

transmitted, which increases (i) the probability of collision and (ii) the channel-accessing waiting-time, hence, harming the video-streaming service. Regarding the impact of the transmission channel conditions, observe that there is an overall decrease on the attained QoE in hostile scenarios (Nakagami- m conditions, Fig. 5.9 (b) and Fig. 5.10 (b)) compared to that obtained in Free Space situations (Fig. 5.9 (a) and Fig. 5.10 (a)). This can be attributed to the impact of the fading channels on the transmitted

packets. The channel variations increase the Bit Error Rate (BER), generating more corrupted packets. Consequently, more retransmissions are necessary, provoking additional collisions and increasing the PLR. Focusing on nodes' motion conditions, it is not noticeable an important decrease in the attained QoE (MOS), but there is a logical increase on the battery drained by the nodes (please, compare Fig. 5.9 (a) and Fig. 5.9 (b) against Fig. 5.10 (a) and Fig. 5.10 (b)). Due to the continuous changes in the routes between transmitters and receivers, the routing tasks become more intense and hence the energy draining, too.

Regarding the timer-based candidate coordination method (in the following JOKER-timer), Fig. 5.11 and 5.12 presents a similar comparison than that discussed above for the case of JOKER-ACK. Note that in these experiments the waiting timer (t_{wait}) has been fixed to 50 ms. Unlike the JOKER-ACK case, now there is not an evident superiority of a specific number of candidates. Although it is clear that the worst results are obtained when 4 candidates are considered, the highest quality levels are attained when employing 2 or 3 candidates depending on the scenario under study. Under heavy traffic conditions, i.e., with 5 simultaneous video streams, the best results are obtained with the lowest number of candidates, 2. In turn, when just 1 video connection is established, the greatest levels of QoE (MOS) are obtained when considering 3 candidates as potential forwarders. This behavior is attributed to the fact that as the number of candidates increases, the probability that one of them is not able to overhear the highest priority candidate also rises. Consequently, more duplicated transmissions happen, harming the overall performance of the network.

Focusing on the impact of the wireless transmission channel conditions, in this case the protocol seems to be more robust than in the JOKER-ACK case. This can be explained by the lack of control-packets exchange in JOKER-timer. First, the effect of fading channels can affect to JOKER-ACK's control messages, provoking more retransmissions, consequently introducing extra delay in the data packet forwarding. Additionally, in the case of JOKER-timer, once the data packet is correctly received by any of the candidates, it will be certainly forwarded after a given time-interval (depending on the candidate priority). This forwarding operation is done without the need of waiting for a confirmation message, which is also prone to be affected by the channel conditions.



(a) Free Space conditions

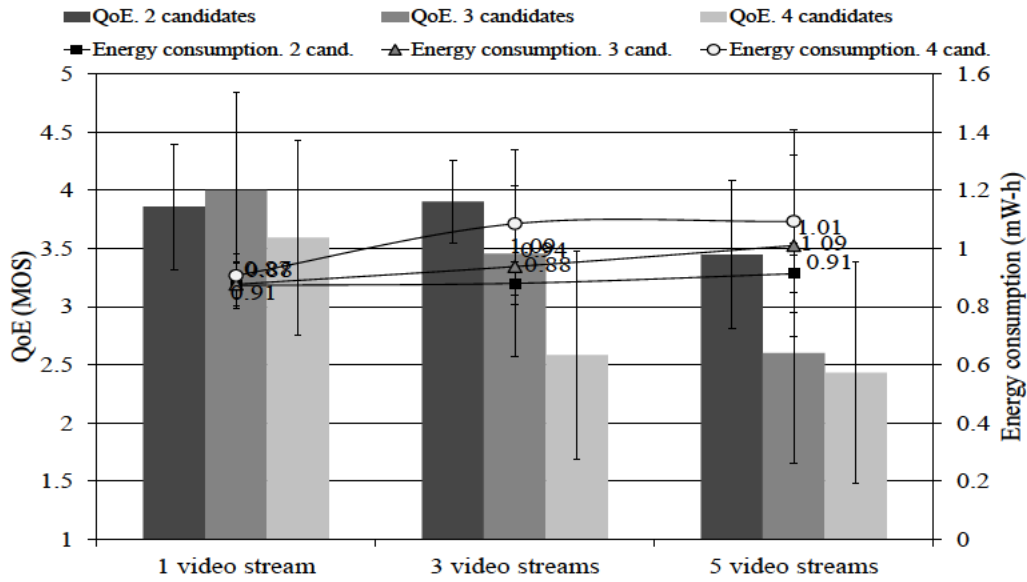
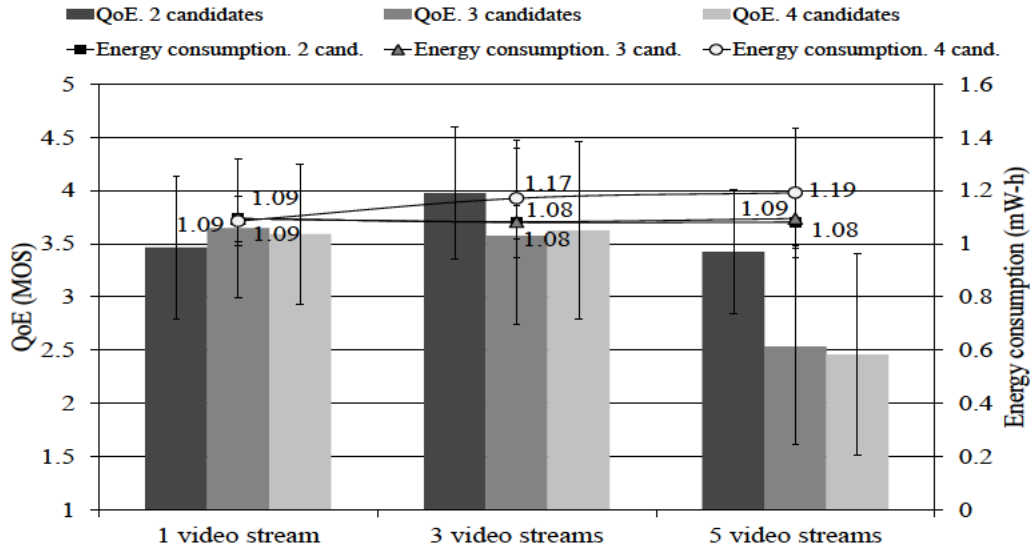

 (b) Nakagami- m conditions

Figure 5.11. Comparison of QoE (MOS) and energy consumption attained by JOKER-timer for a variable number of simultaneous video streams and different number of opportunistic candidates. Results attained in static conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

Comparing the performance of both candidate coordination methodologies, observe the overall improvement in terms of QoE of JOKER-timer (Fig. 5.11 and Fig. 5.12) against JOKER-ACK (Fig. 5.9 and Fig. 5.10). The simplicity of the former adapts better to the needs of the service under consideration. Video-streaming services require quick routing operations, capable of delivering the video packets as soon as possible to the end-extreme of the communication. Although JOKER-ACK ensures the hop-by-hop



(a) Free Space conditions

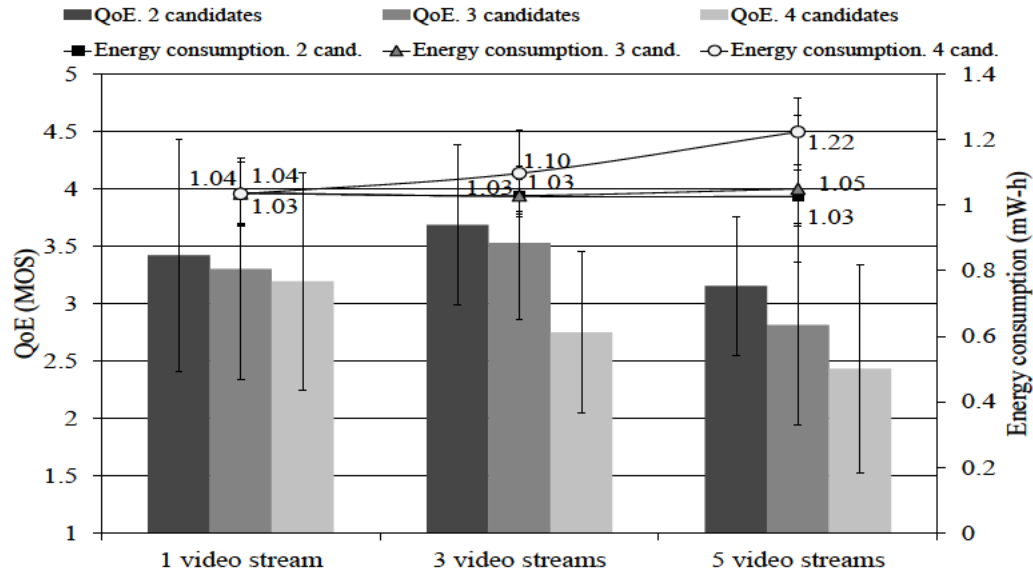

 (b) Nakagami- m conditions

Figure 5.12. Comparison of QoE (MOS) and energy consumption attained by JOKER-timer for a variable number of simultaneous video streams and different number of opportunistic candidates. Results attained in dynamic conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

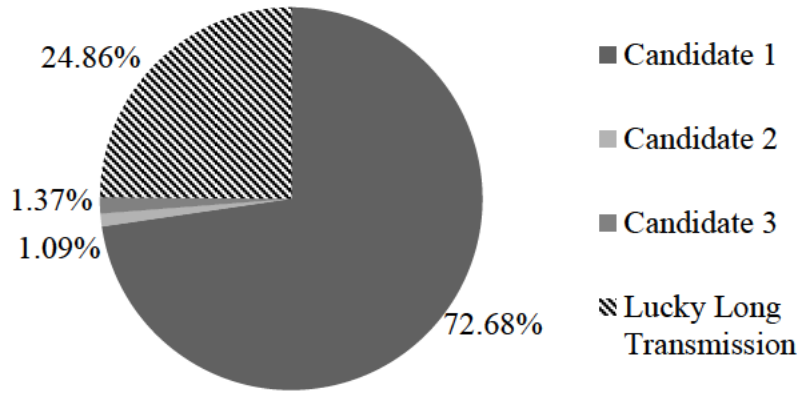
data transmission and no duplicated packets at the receiver, the introduced overhead harms the real-time communication in terms of extra delay and higher collision probability. Besides, the energy consumed by JOKER-timer is also lower than that drained by JOKER-ACK (please compare Fig. 5.9 and Fig. 5.10 with Fig. 5.11 and Fig. 5.12). This is because of the lack of control messages exchange in JOKER-timer that permits the wireless-cards working in low-consumption mode for longer time than in the case of JOKER-ACK.

TABLE 5.5. QoE (MOS) for a variable number of video-streams. Results for JOKER-timer configured with different t_{wait} values

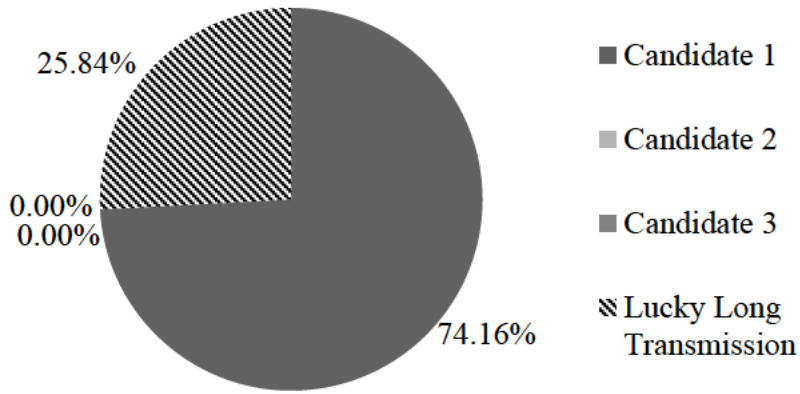
Video-streams	Static				Dynamic			
	25 ms	50 ms	75 ms	100 ms	25 ms	50 ms	75 ms	100 ms
1	3.9	3.9	3.9	3.8	3.3	3.4	3.2	3.4
3	3.3	3.9	3.8	3.3	3.4	3.7	3.2	3.4
5	2.8	3.4	3	2.9	3	3.2	2.8	2.7

Another important configurable parameter in JOKER-timer is the time that one candidate overhears the channel before forwarding a packet (waiting timer, t_{wait}). An excessively short waiting timer could cause the multiple forwarding of the same data packet reducing the network efficiency. In turn, assigning longer values to this parameter introduces extra delay in the case that the candidate with the highest priority fails to forward the packet. Thus, Table 5.5 depicts the reached QoE (MOS) when tuning the t_{wait} parameter for a variable number of simultaneous video streams in the system. The number of considered candidates was set to 2 and the considered environment is that characterized by the realistic Nakagami- m propagation model. Observe how the best results in terms of QoE (MOS) are achieved by setting the waiting timer to 50 ms. This figure is an usual timer employed to configure timer-based candidate coordination schemes [290]. For that reason, in the following the waiting timer is fixed to this value.

In order to explore the actual functionality of the protocol, another experiment aiming at studying the rate of transmissions that take advantage of alternative candidates or lucky long transmissions has been also conducted. In this case, the static topology showed in Fig. 5.1 has been considered. In this experiment, a video transmission similar to that employed so far was established between node TX and node RX and no-artificial link failures were inserted. The Nakagami- m ($m = 5$) propagation model was used to characterize the transmission channel. Fig. 5.13 presents the frequency with which each candidate served as the last hop to the receiver. The lucky long transmissions are also represented. Thereby, observe that for both candidate coordination schemes (JOKER-ACK, Fig 5.13 (a) and JOKER-timer, Fig. 5.13 (b)), most packets are forwarded by the first candidate selected by the TX node.



(a) JOKER-ACK



(b) JOKER-timer

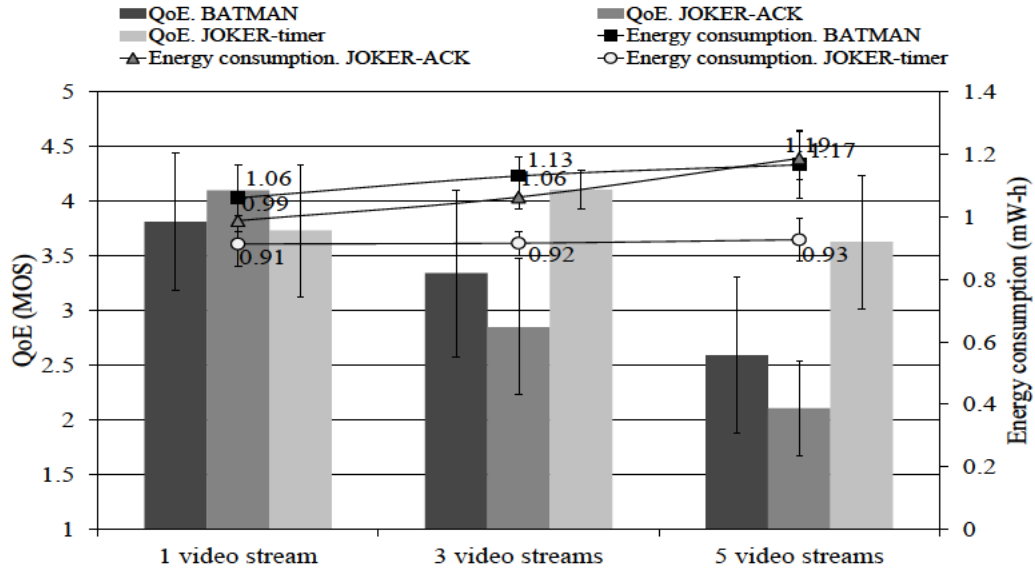
Figure 5.13. Ratio of packets delivered by each candidate to the final destination (see Fig. 5.1). Results attained for JOKER-ACK (a) and JOKER-timer (b) by considering the Nakagami- m propagation model.

This is especially noticeable in the case of JOKER-timer, in which no packets were forwarded by the second and third candidates. The little percentage of packets forwarded by these candidates in JOKER-ACK can be explained by transient saturation periods of the first candidate, which introduces extra delay in the ACK-packet transmission. When this happens, the other candidates' ACK messages are received firstly, so one of them is actually elected to be the packet forwarder. Besides, it is really interesting the high percentage of packets received by lucky long transmissions, i.e., directly from the TX node. In both cases this value is around 25%. By using any other traditional routing algorithm, these packets would be wrongly discarded at the RX node. Besides, extra delay would be added to the communication due to the processing time of the relaying nodes before forwarding the packet.

5.6.2.2 JOKER vs. BATMAN

Once the two candidate coordination schemes have been studied and the best protocol configuration depending on the scenario conditions has been examined, in the following an exhaustive comparison between JOKER and BATMAN is provided. To this end, the number of potential opportunistic candidates has been fixed to 2. As mentioned above, the JOKER-timer's t_{wait} interval has been set to 50 ms. In turn, BATMAN's configuration parameters has been set to their recommended values [204]. Thus, Fig. 5.14 and 5.15 presents a comparison of the QoE and energy consumption obtained for BATMAN, JOKER-ACK, and JOKER-timer in the scenarios previously described. Observe the superior quality values reached by JOKER, especially JOKER-timer, in comparison with those achieved by BATMAN. Focusing in the latter, it is interesting noting how the network performance (in terms of QoE) decreases when the traffic load increases (from 1 to 5 simultaneous video streams). This behavior is similar in the case of JOKER-ACK. In turn, JOKER-timer shows a greater strength supporting high traffic load. Regarding the impact of motion and the wireless channel hostility, observe how JOKER-timer seems to be less affected by these environmental difficulties compared with BATMAN and JOKER-ACK (observe the different scenario conditions in Fig. 5.14 and Fig. 5.15). It is highly remarkable that, using JOKER-timer, the QoE (MOS) level always remains over the notable value of 3, which is a recommended value by the ITU-T for ensuring an acceptable QoE for the end-users.

Focusing on the energy consumed by the wireless cards, again JOKER-timer overcomes the other options. Even in some scenarios, JOKER-ACK also presents lower battery consumption than BATMAN. The reasons for this improvement in the protocol energy efficiency are varied. First, an important factor for decreasing the power consumption is the reduction of the control-packets broadcasts. By means of the dynamic CMSI, the pace at which the routing-packets are broadcasted is severely reduced, so that the wireless card can be in sleep mode for longer periods, permitting energy saving. Another key point for the showed reduction of battery draining is the new metric employed by JOKER for selecting the candidate set, which benefits those nodes farther from the transmitter. By using the *Distance_penalty* factor, the distance progress towards the final destination is increased in each hop, so the routes calculated by JOKER tend to be shorter than those calculated by BATMAN. This is related to the efficiency of JOKER calculating routes. Table 5.6 presents the average number of hops



(a) Free Space conditions

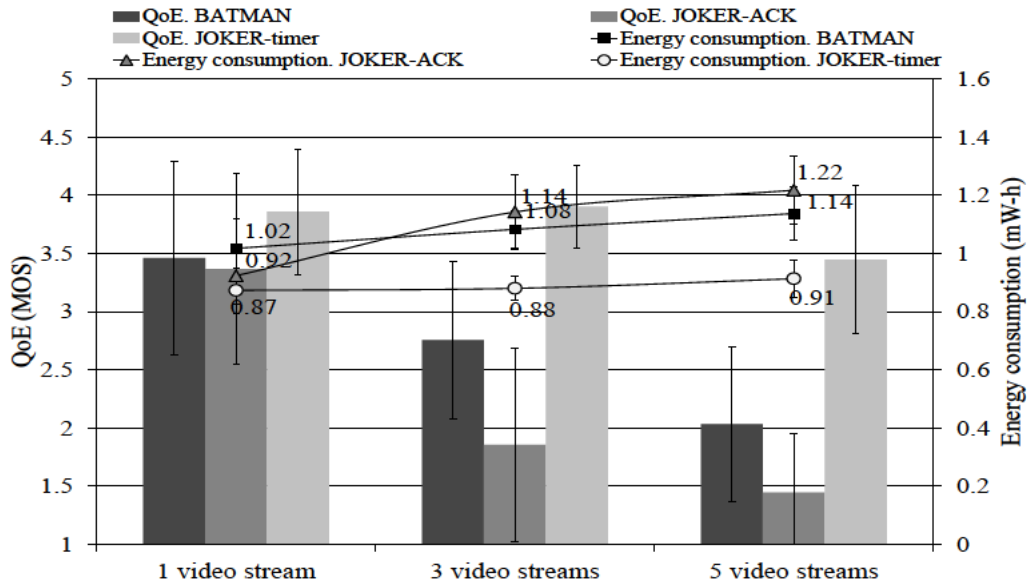

 (b) Nakagami- m conditions

Figure 5.14. Comparison of QoE (MOS) and energy consumption attained by BATMAN, JOKER-ACK, and JOKER-timer for a variable number of simultaneous video streams. Results attained in static conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

composing the path from the transmitter to the final receiver calculated by both BATMAN and JOKER. First, notice that the routes calculated by JOKER-ACK and JOKER-timer have almost the same length. This outcome makes sense as the candidate selection algorithm is not related with the coordination phase. Thus, observe how the routes calculated by JOKER are always shorter than those computed by BATMAN. Besides the new metric employed by JOKER that improves the distance progress in

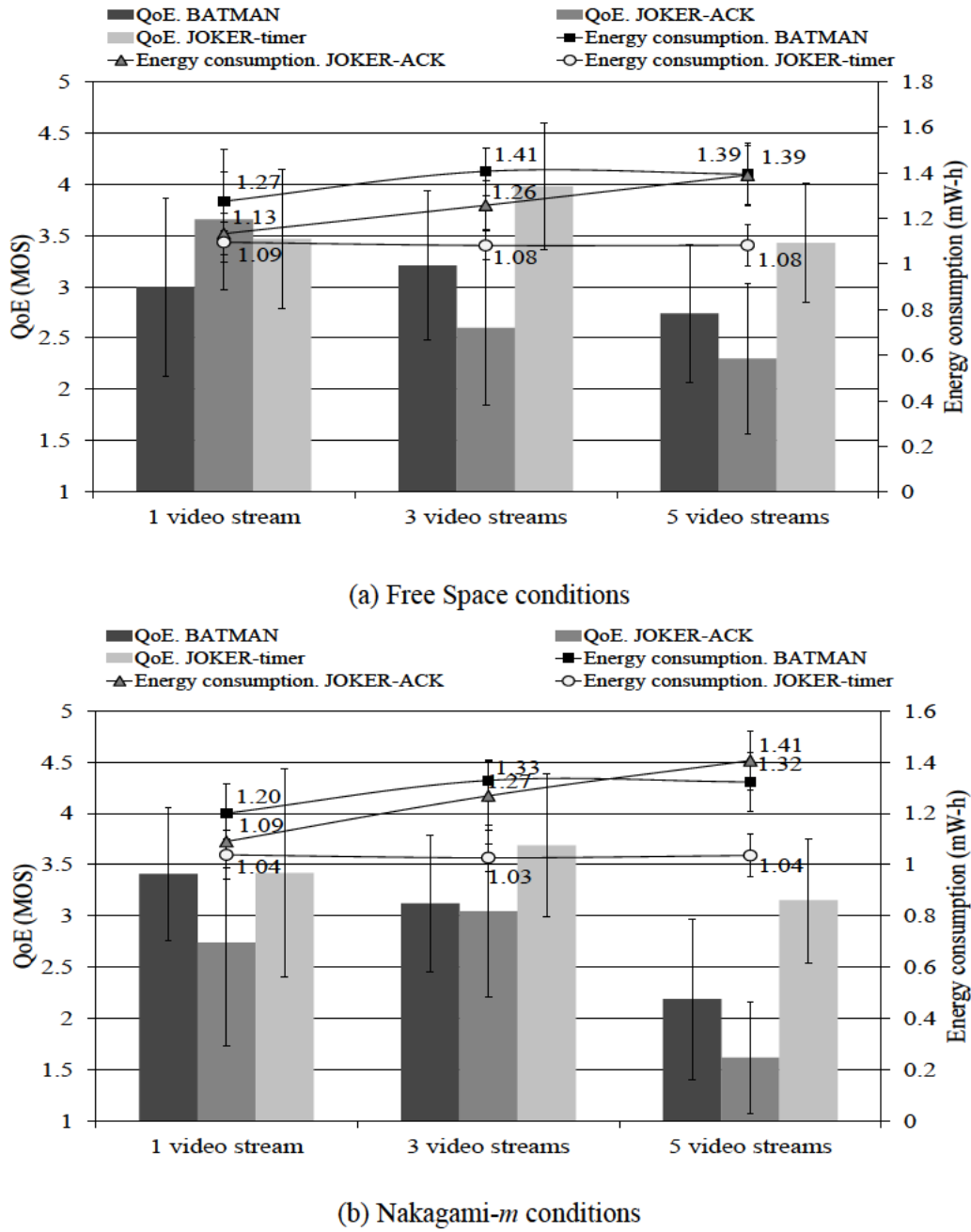


Figure 5.15. Comparison of QoE (MOS) and energy consumption attained by BATMAN, JOKER-ACK, and JOKER-timer for a variable number of simultaneous video streams. Results attained in dynamic conditions by considering the Free Space (a) and Nakagami- m (b) propagations models.

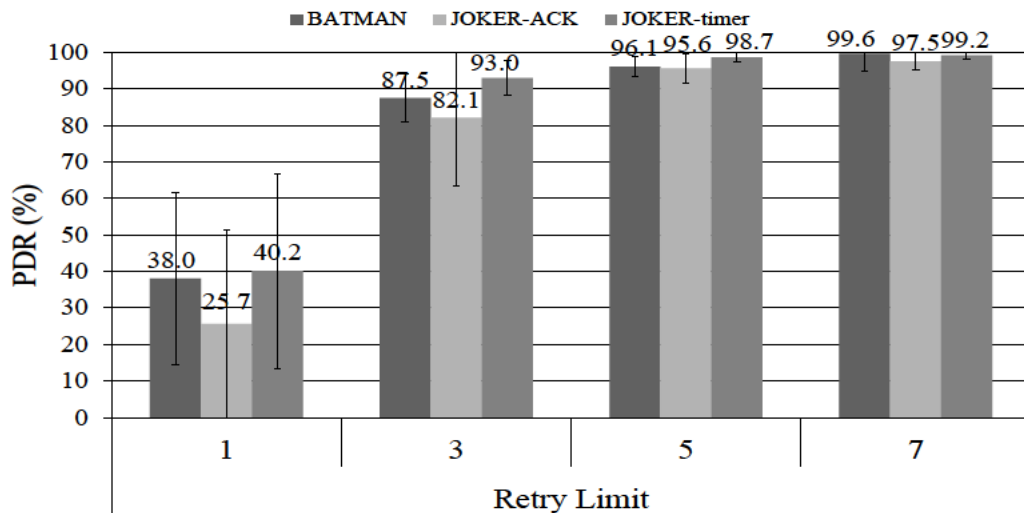
each hop, note that JOKER allows the final destination to take advantage of lucky long transmissions, hence, reducing the number of hops. One surprising result is that the routes calculated in dynamic scenarios are always shorter than those computed in the static ones. This behavior is attributed to the good performance of both protocols; with nodes motion, new opportunities for calculating more efficient routes arise. Thus, the routing algorithms take advantage of this fact, computing more adequate routes than in the static case.

TABLE 5.6. Average hops

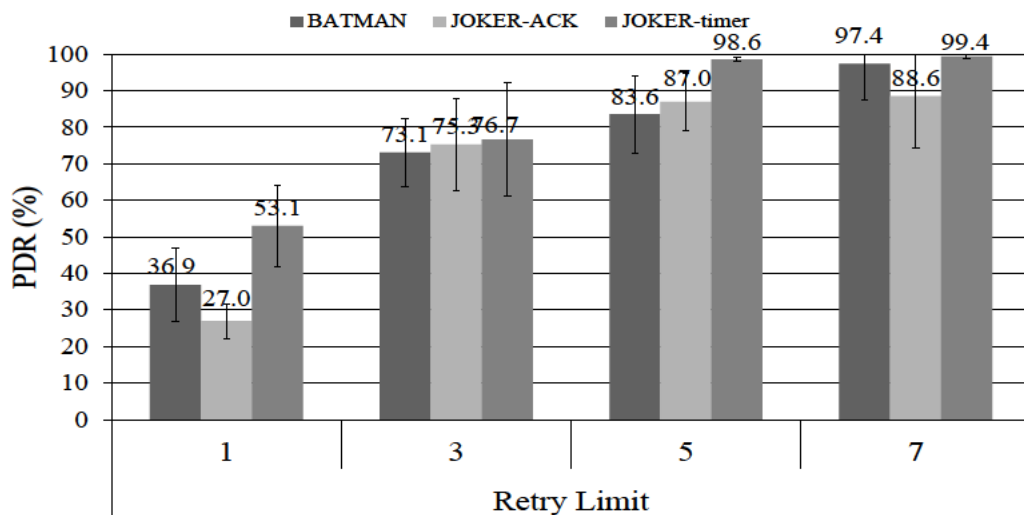
Video-streams	BATMAN		JOKER-ACK		JOKER-timer	
	Static	Dynamic	Static	Dynamic	Static	Dynamic
1	2.38	1.72	2.17	1.26	2.16	1.24
3	2.2	1.55	1.98	1.52	2.04	1.51
5	2.4	1.92	2.37	1.8	2.35	1.8

Another important factor for the network scalability and the node's energy efficiency is the number of retransmissions needed to deliver a packet from the transmitter to the final receiver. The ideal routing protocol would be that one able to successfully deliver a packet to its destination with the lowest number of transmissions. Thereby, Fig. 5.16 depicts the PDR of BATMAN, JOKER-ACK, and JOKER-timer in a static scenario when tuning the IEEE 802.11 retry limit. Note that for the previous experiments, this value was fixed to the standard short retry limit, i.e., 7, which is widely used as default value. It is noticeable the better performance of JOKER-timer, in terms of delivered packets, compared to the other schemes. In turn, BATMAN-ACK presents some difficulties with a low retry limit. The presented results have been extracted from a Nakagami- m hostile scenario; so that JOKER-ACK control packets suffer the impact of the fading channels. If the ACK or Forwarding packets do not finally reach their destinations, the involved data packet is dropped by the forwarder, hence reducing the PDR. Regarding the necessary retransmissions for successfully delivering data packets, observe that reducing the retry limit from 7 down to 5, JOKER-timer reaches PDR levels over the notable figure of 95%. On the other hand, BATMAN still needs 7 retransmissions to reach even lower PDR values. This reduction in the number of necessary retransmissions emphasizes that JOKER-timer presents much more efficient and scalable performance than BATMAN, permitting the improvement of the quality of the streaming service flowing through the network, as well as reducing the energy consumed by the nodes.

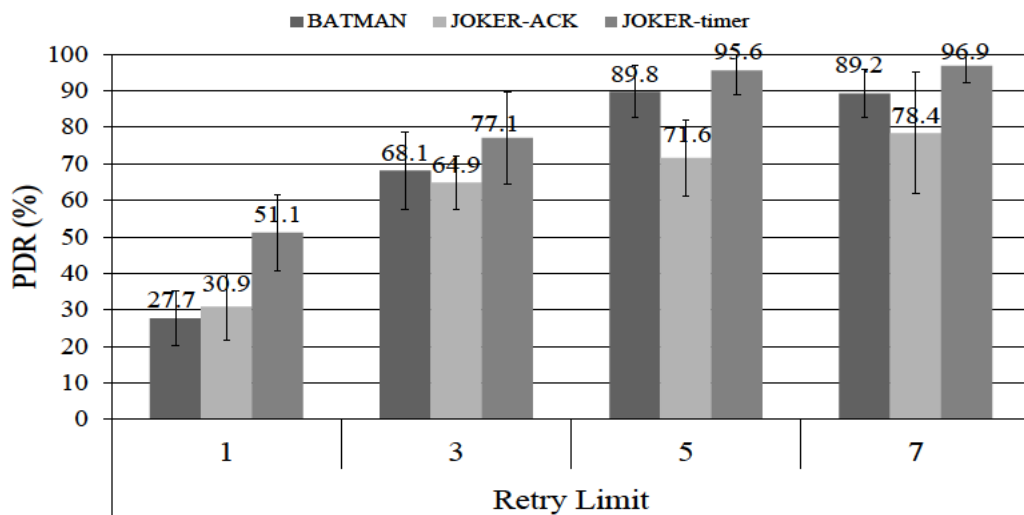
As aforementioned, it has been also conducted an evaluation test by employing the BATMAN and JOKER (3 candidates) implementations for real machines. A similar topology to that showed in Fig. 5.1 was employed. Each test consisted of the transmission of 3 independent 60 s video streams at 150 Kbps between the RX node and



(a) 1 video stream



(b) 3 video streams



(c) 5 video streams

Figure 5.16. PDR attained by BATMAN, JOKER-ACK, and JOKER-timer when tuning the IEEE 802.11 retry limit value. Results from a static scenario characterized by the Nakagami- m propagation model with 1 (a), 3 (b), and 5 (c) simultaneous video streams in the network.

the TX node by means of the relay nodes. Two different network conditions were considered. First, an un-manipulated test was carried out. In this experiment, once the network was correctly configured and the routes well established, the nodes remained untouched. The attained outcomes showed a similar perfect performance by both BATMAN and JOKER protocols, with throughput at the RX node of 150 Kbps and PLR of 0%. These results are not presented here due to their lack of interest. However, in the manipulated test, the performance of both protocols changed. In this test, the relay node's wireless cards were randomly switched off for a time interval of 5 s during the transmission between the TX node and the RX node (please, see the experiment description in section 5.5.2). This behavior was employed in order to simulate link failures provoking disconnection events. Thus, Fig. 5.17 presents the attained throughput and PLR for the 5 performed tests. Observe how both JOKER-ACK and JOKER-timer overcome BATMAN in all the conducted tests in terms of both metrics under consideration. Focusing on the throughput achieved in the RX node (Fig. 5.17 (a)), JOKER-timer presents a minimum value of 135 Kbps, which is a 10% reduction from the original stream. The minimum throughput achieved by JOKER-ACK is 130 Kbps (less than the 15% of reduction from the transmitted flow). In turn, the maximum throughput reached by BATMAN is just 116 Kbps (a 22.6% decrease from the original sent traffic). Regarding PLR, which is a key metric with a deep impact on the quality perceived by the end-users (Fig. 5.17 (b)), the differences between both algorithms are also notable. Observe that JOKER maintains low levels of packet loss (almost of figures around 0% except a maximum value of 4% for JOKER-ACK in test 3); meanwhile BATMAN is not able to obtain values lower than 9%, with several tests achieving a PLR above 20%. The superior performance of JOKER can be attributed to the ready-to-use backup routes that the TX node employs when the first candidate is down. When this happens, the next candidate automatically forwards the packet, so the transmission continues without undesirable cuts. On the other hand, with BATMAN, the TX node needs to re-calculate a new route each time that the selected next-hop towards the RX node is down. Comparing both candidate-selection algorithms, similar to the simulation environment, the best results are attained by JOKER-timer. Nevertheless the performance of JOKER-ACK seems to be sufficient to support video-traffic in the studied topology.

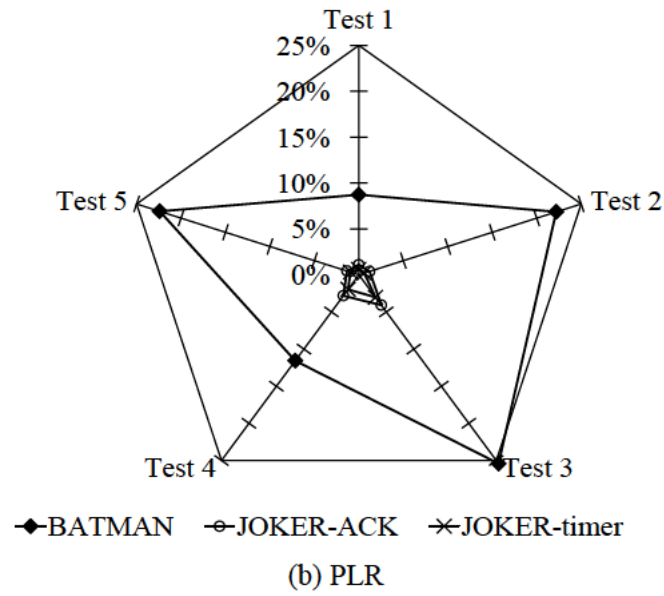
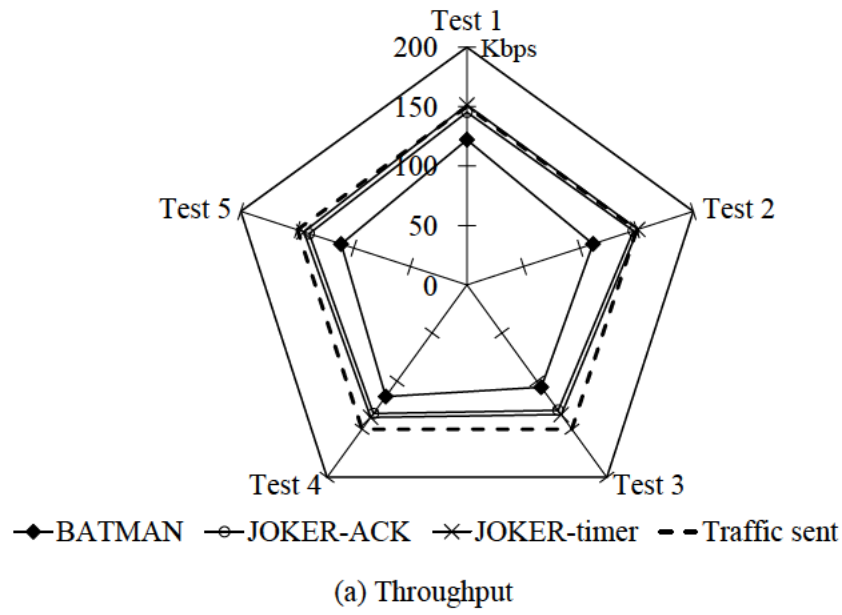


Figure 5.17. Performance comparison of JOKER and BATMAN implementations for real machines. Throughput (a) and PLR (b).

TABLE 5.7. QoE (MOS) values attained in the experimental test-bench

	BATMAN	JOKER-ACK	JOKER-timer
Test 1	1.28	3.46	3.85
Test 2	1	3.39	3.84
Test 3	1	2.05	2.42
Test 4	1.1	2.5	3
Test 5	1	3.3	3.6

From a QoE perspective, the high levels of PLR attained by BATMAN causes a great decrease on the video quality, achieving QoE (MOS) values near 1 in all the conducted tests (please, see Table 5.7). Recall that the QoE estimations in terms of MOS have been computed by employing expression (5.4). In turn, JOKER presents acceptable QoE (MOS) values above 3 in most of the performed tests. Again, the best results are attained by JOKER-timer that achieves remarkable QoE (MOS) values above 3.6 in three out of five tests.

Therefore, it has been demonstrated from both QoS and QoE perspectives the superior performance of JOKER over BATMAN, also in real scenarios. In the lights of these results, it can be concluded that JOKER is an efficient proposal for the distribution of streaming traffic in ad-hoc networks, considering the energy constraints imposed by these systems, too.

5.7 Conclusion

In this chapter it has been discussed about two novel networking paradigms: green networking and opportunistic routing. First it has been presented a comparison study evaluating the energy consumed by three prominent ad-hoc routing protocols (BATMAN, OLSR, and AODV) supporting a video-streaming service under different environmental conditions. Besides, the quality attained by the multimedia service in terms of QoE has been also considered. Thus, the trade-off between the quality perceived when consuming the video-service and the energy consumed by the end-device has been analyzed. From this study, it can be concluded that, on the one hand, in environments with low level of packet loss, i.e., Free Space situations, there is not a clear trade-off between the QoE attained by the multimedia communication and the needed energy for routing tasks. On the other hand, under lossy conditions, e.g., Nakagami- m environments, the best results in terms of QoE are achieved by the protocol that wastes more energy in their routing operations, i.e., BATMAN.

Then, it has been presented an opportunistic routing protocol, so-called JOKER, addressing the aforementioned trade-off between QoE in multimedia transmissions and energy consumption. Following the opportunistic paradigm, JOKER presents novelties in both the candidate selection and coordination phases. Focusing on the former, a new metric has been introduced, gathering the packet-delivery reliability of the links with the distance-progress towards the final destination. Additionally, a dynamical adaptation of

the protocol's control-message sending-interval was developed aiming at adapting JOKER to the actual network conditions and reducing energy consumption as well. This enhancement, which was initially tested in BATMAN, makes this algorithm a more battery-friendly protocol, even increasing the quality of the considered video service. Regarding candidate coordination, two different procedures were included in JOKER, namely, ACK-based and timer-based coordinations. Two different JOKER implementations, for simulation and experimental test-benches, were introduced and tested, allowing comparing the performance of JOKER supporting video-streaming traffic with that presented by the prominent ad-hoc routing algorithm BATMAN. From the attained results, it was extracted the superior performance of JOKER in terms of both QoE and energy efficiency in comparison with BATMAN. These outcomes were confirmed in the several scenarios evaluated in the different-nature test-benches under consideration. In the lights of the presented results, it can be concluded that JOKER is an efficient proposal for the distribution of streaming traffic in ad-hoc networks, considering the energy constraints imposed by these systems, too.

5.8 Contributions Related to this Chapter

R. Sanchez-Iborra and M.-D. Cano, “Evaluation of the trade-off between energy consumption and QoE in video transmission with MANET routing protocols,” in *XII Jornadas de Ingenieria Telematica (JITEL 2015)*, 2015, pp. 145 – 152.

R. Sanchez-Iborra and M.-D. Cano, “JOKER: a novel opportunistic routing protocol,” *IEEE Journal on Selected Areas in Communications*. Under review. (I. F.: 4.138).

Chapter 6

Conclusion

6.1 Conclusion

Within the next 5 years, monthly global mobile data traffic will surpass 24.3 Exabytes, and nearly three-fourths of the world's mobile data traffic will be multimedia content [291]. In all of its forms (VoIP, video on demand, live streaming, conferencing, etc.), multimedia traffic is one of the most demanding services in terms of network efficiency and reliability, being for instance extremely low tolerant to packet loss and delay variations. Therefore, fulfilling a high-quality multimedia service becomes a must for service providers and network operators worldwide. This is especially challenging to achieve in wireless networks. These systems are, certainly, those which provide the most attractive features to customers, such as mobility, or configuration and deployment simplicity. However, reaching acceptable levels of quality in wireless environments is much tougher than in wired networks. Wireless systems present new impairments that make them unpredictable and difficult to keep under perfect control. Concretely, these impairments are related to the variability of the wireless transmission channel: interferences, fading, shadowing, etc. can severely degrade wireless communications.

For those reasons, the main focus of this thesis has been evaluating and understanding the multimedia-content delivery-process in wireless networks, aiming at presenting different proposals to improve the quality reached by these services. In the following, the results achieved on each of the previous chapters are summarized:

Chapter 2. In this chapter it was presented a global overview of the QoE concept, dissecting its different dimensions and discussing about this paradigm from several angles. This chapter provides the necessary background to understand the QoE-perspective given to the rest of the dissertation. Aiming at supporting this discussion, two exhaustive reviews addressing the most relevant methodologies for estimating multimedia QoE by means of objective techniques were presented. These surveys were focused on VoIP and Future TV services quality estimation algorithms, exploring the best suited models depending on the scenario and service under evaluation. Finally, aiming at implementing in a real experiment all the concepts examined so far, a real

VoIP QoE subjective test was described. The results, attained from both subjective and objective perspectives, demonstrated the relationship between these approaches. One of the main findings extracted from this study is that the impact on the QoE of other impairments different from the usually evaluated, i.e., PLR and delay, is particularly relevant in scenarios with low level of network-introduced impairments.

Chapter 3. The performance of Wi-Fi networks configured in infrastructure mode supporting traffic with stringent requirements such as VoIP was studied. To this end, a realistic study from QoE perspective evaluating the impact of the transmission channel conditions on the quality of the VoIP service was thoroughly presented. Besides, the negative effects of the fading channels on the network's capacity and coverage range were also studied. Finally, a comparison between the effect of different voice coding schemes (no-compression codec vs. low bit-rate codec) and packetization intervals on the voice calls' quality and the system's capacity was conducted. From the obtained results, it can be extracted the need of characterizing the wireless physical layer as realistic as possible by means of proper propagation models. Ignoring the effects introduced by the transmission medium could lead to inaccurate results, especially in studies based on simulation. Regarding the impact of the coding-scheme, it was also showed that the use of low bit-rate codecs allows an increase on the system the capacity and coverage range but at the expense of decreasing the QoE in comparison with no-compression codecs. Finally, the impact of setting different packetization time-intervals was also evaluated, obtaining advantageous results for the packetization of 20 ms because, without decreasing the level of quality, the network capacity almost doubles compared with the 10 ms packetization. To sum up, it was concluded that, by means of a good systems configuration, Wi-Fi networks in infrastructure mode are capable to support highly demanding traffic such as VoIP communications with good levels of quality.

Chapter 4. The capability of WiFi networks in ad-hoc mode (Mobile Ad-hoc Network, MANET) to support multimedia traffic (VoIP and video-streaming) was explored. Thus, the prominent proactive ad-hoc routing protocols BATMAN and OLSR were evaluated in two different-nature test-benches, namely, a realistic scenario and a simulation framework. This variety of test-beds permitted analyzing several scenarios regarding the traffic pattern employed, the area covered by the network, the number of nodes, or their mobility characteristics, among others. A statistical comparison of the

results attained in both test-benches that allowed validating the outcomes from the simulation framework was performed. Finally, it was presented some adjustments to the ad-hoc routing protocols that permit the improvement of the multimedia services' QoE. Concretely, the impact of tuning the BATMAN OGM interval and the OLSR HELLO interval on the quality achieved by the multimedia service flowing through the ad-hoc network was evaluated. From the obtained results, it can be concluded that by using the mentioned time-intervals tuning, some possible enhancements to the BATMAN routing protocol aiming at improving the QoE of the multimedia service flowing through the network are feasible. This interval-tuning concept opens a new way to follow for improving the overall performance of the routing protocol and, consequently, of the multimedia service transmitted over it.

Chapter 5. In this chapter, two novel networking paradigms were considered: green networking and opportunistic routing. First, an exhaustive evaluation of the trade-off between energy consumption and QoE in video transmission with different MANET routing protocols was conducted. From this study, it was concluded that in environments with low level of packet loss, there is not a clear trade-off between the QoE attained by the multimedia communication and the needed energy for routing tasks. On the other hand, in more hostile scenarios the best results in terms of QoE were achieved by the protocol that wastes more energy in their routing operations. Thereafter, the ad-hoc routing concept was explored from a different perspective: opportunistic networks. Concretely, a novel opportunistic routing protocol, so-called JOKER was presented. Within this protocol, a new metric for selecting the best next-hop towards the final destination was developed. This new metric ensures high packet-delivery reliability, increasing in each hop the distance-progress towards the final destination. A dynamical adaptation of the control-message sending-interval was also developed. This enhancement, which was initially tested in BATMAN, further improves the capability of this protocol to support highly demanding services, such as video-streaming applications; additionally, it permits reducing the energy consumed by the nodes in routing tasks. Finally, two different JOKER implementations, for simulation and experimental test-benches, were introduced and tested, allowing comparing the performance of JOKER supporting video-streaming traffic with that presented by BATMAN. The attained results showed, from a QoE perspective, the superior performance of JOKER supporting multimedia traffic in comparison with BATMAN.

The new proposal also showed a greater energy-efficiency, reducing the power consumed by the nodes in routing tasks.

References

- [1] Cisco, “Cisco visual networking index: forecast and methodology, 2013–2018. White paper,” *White Pap.*, 2014.
- [2] “Methods for subjective determination of transmission quality,” *ITU-T Recomm. P.800*, 1996.
- [3] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, and G. Pujolle, “Quality of Experience of VoIP service: a survey of assessment approaches and open issues,” *IEEE Commun. Surv. Tutorials*, vol. 14, no. 2, pp. 491–513, 2012.
- [4] Y. Chen, K. Wu, and Q. Zhang, “From QoS to QoE: a survey and tutorial on state of art, evolution and future directions of video quality analysis,” *IEEE Commun. Surv. Tutorials*, vol. PP, no. 99, p. 1, 2014.
- [5] M. Conti and S. Giordano, “Multihop Ad Hoc Networking: The Evolutionary Path,” in *Mobile Ad Hoc Networking*, John Wiley & Sons, Inc., 2013, pp. 1–33.
- [6] P. Bellavista, A. Corradi, and L. Foschini, “Self-organizing seamless multimedia streaming in dense Manets,” *IEEE Pervasive Comput.*, vol. 12, no. 1, pp. 68–78, 2013.
- [7] T. Vanhatupa, M. Hainnikainen, and T. D. Hdmalainen, “Multihop IEEE 802.11 b WLAN Performance for VoIP,” in *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC’ 05)*, 2005, pp. 1925–1930.
- [8] C.-O. Chow and H. Ishii, “Enhancing real-time video streaming over mobile ad hoc networks using multipoint-to-point communication,” *Comput. Commun.*, vol. 30, no. 8, pp. 1754–1764, Jun. 2007.
- [9] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A. Joglekar, “An integrated experimental environment for distributed systems and networks,” *ACM SIGOPS Oper. Syst. Rev. - OSDI ’02 Proc. 5th Symp. Oper. Syst. Des. Implement.*, vol. 36, no. SI, pp. 255 – 270, 2002.
- [10] “Omnet++.” [Online]. Available: <<http://www.omnetpp.org>>.
- [11] B. E. Carpenter and K. Nichols, “Differentiated services in the Internet,” *Proc. IEEE*, vol. 90, no. 9, pp. 1479–1494, Sep. 2002.
- [12] M.-D. Cano and F. Cerdan, “Subjective QoE analysis of VoIP applications in a wireless campus environment,” *Telecommun. Syst.*, vol. 49, no. 1, pp. 5–15, Jun. 2010.
- [13] “Definitions of terms related to quality of service,” *ITU-T Recomm. E.800*, 2008.

- [14] “Communications Quality of Service: a framework and definitions,” *ITU-T Recomm. G.1000*, 2001.
- [15] “New definitions for inclusion in Recommendation ITU-T P.10/G.100,” *ITU-T Recomm. P.10/G.100 Amend. 2*, 2008.
- [16] “End-user multimedia QoS categories,” *ITU-T Recomm. G.1010*, 2001.
- [17] “Internet protocol data communication service – IP packet transfer and availability performance parameters,” *ITU-T Recomm. Y.1540*, 2011.
- [18] “Network performance objectives for IP-based services,” *ITU-T Recomm. Y.1541*, 2011.
- [19] K. Rehman Laghari and K. Connelly, “Toward total quality of experience: a QoE model in a communication ecosystem,” *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 58–65, Apr. 2012.
- [20] W. Song and D. W. Tjondronegoro, “Acceptability-based QoE models for mobile video,” *IEEE Trans. Multimed.*, vol. 16, no. 3, pp. 738–750, Apr. 2014.
- [21] D. H. Shin, “Determinants of customer acceptance of multi-service network: An implication for IP-based technologies,” *Inf. Manag.*, vol. 46, no. 1, pp. 16–22, Jan. 2009.
- [22] M. Volk, J. Sterle, U. Sedlar, and A. Kos, “An approach to modeling and control of QoE in next generation networks,” *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 126–135, Aug. 2010.
- [23] J. Lassalle, L. Gros, T. Morineau, and G. Coppin, “Impact of the content on subjective evaluation of audiovisual quality: What dimensions influence our perception?,” in *IEEE international Symposium on Broadband Multimedia Systems and Broadcasting*, 2012, pp. 1–6.
- [24] “Requirements for the support of IPTV services,” *ITU-T Recomm. Y.1901*, 2009.
- [25] “Fundamental voice transmission objectives for VoIP terminals and gateways,” *ITU-T Recomm. P.1010*, 2007.
- [26] TelcoBridges, “Top 5 VoIP customer complaints.”
- [27] “Perceptual evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” *ITU-T Recomm. P.862*, 2001.
- [28] “Perceptual objective listening quality assessment,” *ITU-T Recomm. P.863*, 2014.
- [29] W. Yang, “Enhanced Modified Bark Spectral Distortion (EMBSD): an objective speech quality measure based on audible distortion and cognition model,” 1999.

-
- [30] T. Lee and J. Pan, "Improving R-score of adaptive VoIP codec in IEEE 802.16 networks," in *14th AsiaPacific Conference on Communications*, 2008, pp. 1 – 5.
 - [31] S. Hoh and D. Singh, "Effectiveness of adaptive codec switching VoIP application over heterogeneous networks," in *2nd Asia Pacific Conference on Mobile Technology, Applications and Systems*, 2005.
 - [32] "Single-ended method for objective speech quality assessment in narrow-band telephony applications," *ITU-T Recomm. P.563*, 2004.
 - [33] "The E-model: a computational model for use in transmission planning," *ITU-T Recomm. G.107*, 2011.
 - [34] "Conformance testing for voice over IP transmission quality assessment models," *ITU-T Recomm. P.564*, 2007.
 - [35] A. W. Rix, J. G. Beerends, D.-S. Kim, P. Kroon, and O. Ghitza, "Objective assessment of speech and audio quality - technology and applications," *Audio, Speech, Lang. Process. IEEE Trans.*, vol. 14, no. 6, pp. 1890–1901, 2006.
 - [36] S. Moller, W.-Y. Chan, N. Cote, T. H. Falk, A. Raake, and M. Waltermann, "Speech quality estimation: models and trends," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 18–28, 2011.
 - [37] A. E. Mahdi and D. Picovici, "Advances in voice quality measurement in modern telecommunications," *Digit. Signal Process.*, vol. 19, no. 1, pp. 79–103, Jan. 2009.
 - [38] "In-service non-intrusive measurement device - Voice service measurements," *ITU-T Recomm. P.561*, 2002.
 - [39] "Definition of categories of speech transmission quality," *ITU-T Recomm. G.109*, 1999.
 - [40] "Wideband E-model," *ITU-T Recomm. G.107.1*, 2011.
 - [41] S. Möller, A. Raake, N. Kitawaki, and A. Takahashi, "Impairment factor framework for wide-band speech codecs," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 14, no. 6, pp. 1969–1976, 2006.
 - [42] "Transmission impairments due to speech processing," *ITU-T Recomm. G.113*, 2007.
 - [43] M. Wältermann, A. Raake, and S. Möller, "Extension of the E-model towards super-wideband speech transmission," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP'10)*, 2010, no. 1, pp. 4654–4657.

- [44] A. E. Conway, "Output-based method of applying PESQ to measure the perceptual quality of framed speech signals," in *IEEE Wireless Communications and Networking Conference (WCNC'04)*, 2004, vol. 4, pp. 2521–2526 Vol.4.
- [45] A. D. Clark, "Modeling the effects of burst packet loss and recency on subjective voice quality," in *IP Telephony Workshop*, 2001.
- [46] J. H. Rosenbluth, "Testing the quality of connections having time varying impairments," *ITU-T del. cont. COM12-D64*, 1998.
- [47] P. Gray, R. Massara, and M. Hollier, "An experimental investigation of the accumulation of perceived error in time-varying speech distortions," 1997.
- [48] L. Carvalho, E. Mota, R. Aguiar, A. F. Lima, J. N. de Souza, and A. Barreto, "An E-model implementation for speech quality evaluation in VoIP systems," in *10th IEEE Symposium on Computers and Communications (ISCC'05)*, 2005, pp. 933–938.
- [49] R. G. Cole and J. H. Rosenbluth, "Voice over IP performance monitoring," *Comput. Commun. Rev.*, vol. 31, no. 2, pp. 9–24, 2001.
- [50] L. Ding and R. A. Goubran, "Speech quality prediction in VoIP using the extended E-model," in *Global Telecommunications Conference, GLOBECOM'03*, 2003, vol. 7, pp. 3974 – 3978.
- [51] J. Ren, C. Zhang, W. Huang, and D. Mao, "Enhancement to E-model on standard deviation of packet delay," in *The 3rd International Conference on Information Sciences and Interaction Sciences*, 2010, pp. 256–259.
- [52] H. Z. H. Zhang, L. X. L. Xie, J. B. J. Byun, P. Flynn, and C. S. C. Shim, "Packet loss burstiness and enhancement to the E-model," in *6th Int. Conference on Software Engineering Artificial Intelligence Networking and Parallel Distributed Computing and First ACIS Int. Workshop on Self Assembling Wireless Network*, 2005, pp. 214 – 219.
- [53] M. Grah and P. Radcliffe, "Dynamic QoS and network control for commercial VoIP systems in future heterogeneous networks," in *Tenth IEEE International Symposium on Multimedia*, 2008, pp. 356–363.
- [54] A. Meddahi and H. Afifi, "'Packet-E-model': E-model for VoIP quality evaluation," *Comput. Networks*, vol. 50, no. 15, pp. 2659–2675, Oct. 2006.
- [55] B. Lewcio and M. Waltermann, "E-model supported switching between narrowband and wideband speech quality," in *International Workshop on Quality of Multimedia Experience, QoMEx*, 2009, pp. 98–103.
- [56] W. Li, F. Yang, and S. Wan, "Quality assessment for networked speech based on active packet detection," in *International Conference On Computer and Communication Technologies in Agriculture Engineering (CCTAE)*, 2010, vol. 3, pp. 561–564.

-
- [57] T. Falk and W.-Y. Chan, "Performance study of objective speech quality measurement for modern wireless-VoIP communications," *EURASIP J. Audio, Speech, Music Process.*, vol. 2009, pp. 1–11, 2009.
- [58] D. Picovici and J. Nelson, "Time-varying quality estimation for VoIP over Wireless Networks," in *9th IFIP International Conference on Mobile Wireless Communications Networks*, 2007, pp. 91–95.
- [59] S. Jelassi and H. Youssef, "Connectivity Aware Instrumental Approach for Measuring Vocal Transmission Quality Over a Wireless Ad-Hoc Network," in *New Technologies, Mobility and Security*, 2008, pp. 1–5.
- [60] S. Jelassi, H. Youssef, L. Sun, and G. Pujolle, "NIDA: a parametric vocal quality assessment algorithm over transient connections," in *12th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services: Wired-Wireless Multimedia Networks and Services Management*, 2009, pp. 106–117.
- [61] A. Passito, E. Mota, and R. Aguiar, "Using an E-model implementation to evaluate speech quality in voice over 802.11 b networks with VPN/IPSec," in *International Conference on Wireless Communications, Networking and Mobile Computing*, 2005, vol. 2, pp. 1123–1127.
- [62] A. Passito and A. Nascimento, "Management of VoWLAN security parameters using a QoS tool," in *Wireless Communications and Networking Conference. WCNC'06*, 2006, vol. 2, pp. 611–616.
- [63] "Analysis and interpretation of INMD voice-service measurements," *ITU-T Recomm. P.562*, 2004.
- [64] Psytechnics Inc., "CCI: Getting the message loud and clear...measuring the clarity of speech over networks," *CCI White Pap.*, 2001.
- [65] J. S. Hartanto, "QualiView: an in-service, non-intrusive monitoring device (INMD) for call quality performance monitoring," in *Tenth International Conference on Digital Satellite Communications*, 1995, vol. 2, pp. 682 – 689.
- [66] "SwissQual." [Online]. Available: <www.swissqual.com>.
- [67] "Psytechnics." [Online]. Available: <www.netscout.com>.
- [68] "Opticom." [Online]. Available: <www.opticom.de>.
- [69] D.-S. Kim, "ANIQUE: an auditory model for single-ended speech quality estimation," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 5, pp. 821–831, Sep. 2005.
- [70] D.-S. Kim and A. Tarraf, "ANIQUE+: a new American national standard for non-intrusive estimation of narrowband speech quality," *Bell Labs Tech. J.*, vol. 12, no. 1, pp. 221–236, May 2007.

- [71] L. Malfait, J. Berger, and M. Kastner, "P.563 - The ITU-T standard for single-ended speech quality assessment," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 14, no. 6, pp. 1924–1934, 2006.
- [72] "Subjective performance assessment of telephone-band and wideband digital codecs," *ITU-T Recomm. P.830*, 1996.
- [73] P. Gray, M. P. Hollier, and R. E. Massara, "Non-intrusive speech-quality assessment using vocal-tract models," in *Vision, Image and Signal Processing. IEEE*, 2000, vol. 147, no. 6, pp. 493–501.
- [74] A. Ekman and B. Kleijn, "Improving quality prediction accuracy of P. 563 for noise suppression," in *International Workshop Acoustic Echo and Noise Control*, 2008.
- [75] T. H. Falk, S. Möller, V. Karaiskos, and S. King, "Improving instrumental quality prediction performance for the Blizzard Challenge," in *International TTS Challenge Workshop*, 2008.
- [76] M. Abareghi, M. M. Homayounpour, M. Dehghan, and A. Davoodi, "Improved ITU-P.563 non-intrusive speech quality assessment method for covering VOIP conditions," in *10th International Conference on Advanced Communication Technology*, 2008, pp. 354–357.
- [77] S. Broom and M. Hollier, "Speech quality measurement tools for dynamic network management," *MESAQIN*, 2003.
- [78] "Mapping function for transforming P.862 raw result scores to MOS-LQO," *ITU-T Recomm. P.862.1*, 2003.
- [79] F. Neves, S. Cardeal, S. Soares, P. Assuncao, and F. Tavares, "Quality model for monitoring QoE in VoIP services," *IEEE EUROCON - Int. Conf. Comput. as a Tool*, pp. 1–4, Apr. 2011.
- [80] T. H. Falk and W.-Y. Chan, "Nonintrusive speech quality estimation using Gaussian mixture models," *IEEE Signal Process. Lett.*, vol. 13, no. 2, pp. 108–111, Feb. 2006.
- [81] T. H. Falk and W. Y. Chan, "Enhanced non-intrusive speech quality measurement using degradation models," in *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, 2006, vol. 1, pp. 837–840.
- [82] T. H. Falk, H. Yuan, and W. Y. Chan, "Single-ended quality measurement of noise suppressed speech based on Kullback-Leibler distances," *J. Multimed.*, vol. 2, no. 5, pp. 19–26, Sep. 2007.
- [83] T. H. Falk and W. Y. Chan, "Single-ended speech quality measurement using machine learning methods," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 14, no. 6, pp. 1935–1947, Nov. 2006.

-
- [84] J. Wang, J. Luo, S. Zhao, and J. Kuang, "Non-intrusive objective speech quality measurement based on GMM and SVR for narrowband and wideband speech," in *11th IEEE Singapore International Conference on Communication Systems*, 2008, pp. 193–198.
 - [85] J. Wang, Y. Zhang, Y. Song, S. Zhao, and J. Kuang, "An improved non-intrusive objective speech quality evaluation based on FGMM and FNN," in *3rd International Congress on Image and Signal Processing*, 2010, pp. 3495–3499.
 - [86] L. Sun and E. C. Ifeachor, "Perceived speech quality prediction for voice over IP-based networks," in *IEEE International Conference on Communications. ICC'02*, 2002, vol. 4, pp. 2573–2577.
 - [87] S. Mohamed, G. Rubino, and M. Varela, "Performance evaluation of real-time speech through a packet network: a random neural networks-based approach," *Perform. Eval.*, vol. 57, no. 2, pp. 141–161, Jun. 2004.
 - [88] S. Mohamed, G. Rubino, and M. Varela, "A method for quantitative evaluation of audio quality over packet networks and its comparison with existing techniques," in *Measurement of Speech and Audio Quality in Networks. MESAQIN'04*, 2004.
 - [89] A. P. C. da Silva, M. Varela, E. de Souza e Silva, R. M. M. Leão, and G. Rubino, "Quality assessment of interactive voice applications," *Comput. Networks*, vol. 52, no. 6, pp. 1179–1192, Apr. 2008.
 - [90] W. Cherif, A. Ksentini, D. Negru, and M. Sidibe, "A_PSQA: PESQ-like non-intrusive tool for QoE prediction in VoIP services," in *IEEE International Conference on Communications (ICC)*, 2012, pp. 2124–2128.
 - [91] A. Mahdi and D. Picovici, "New single-ended objective measure for non-intrusive speech quality evaluation," *Signal, Image Video Process.*, vol. 4, pp. 23 – 38, 2010.
 - [92] T. Hoßfeld, P. Tran-Gia, and M. Fiedler, "Quantification of quality of experience for edge-nased applications," in *20th International Teletraffic Congress*, 2007, pp. 361 – 373.
 - [93] T. Hoßfeld, D. Hock, P. Tran-Gia, K. Tutschku, and M. Fiedler, "Testing the IQX hypothesis for exponential interdependency between QoS and QoE of voice codecs iLBC and G.711," in *18th ITC Specialist Seminar on Quality of Experience*, 2008.
 - [94] M. Soloducha and A. Raake, "Speech quality of VoIP: bursty packet loss revisited," in *ITG-Fachbericht 252: Speech Communication*, 2014, pp. 1–4.
 - [95] C. Manzano and Y. Jung, "Burst packet loss and enhanced packet loss-based quality model for mobile voice-over Internet protocol applications," *IET Commun.*, vol. 8, no. 1, pp. 41–49, Jan. 2014.

- [96] A. Raja, R. M. A. Azad, C. Flanagan, and C. Ryan, "Real-time, non-intrusive evaluation of VoIP," *Genet. Program.*, vol. 4445, pp. 217–228, Apr. 2007.
- [97] A. Raja, R. M. A. Azad, C. Flanagan, and C. Ryan, "Evolutionary speech quality estimation in VoIP," *Soft Comput.*, vol. 15, no. 1, pp. 89–94, Nov. 2009.
- [98] S. Jelassi, H. Youssef, C. Hoene, and G. Pujolle, "Voicing-aware parametric speech quality models over VoIP networks," in *Global Information Infrastructure Symposium*, 2009, pp. 1–8.
- [99] L. Ding, Z. Lin, A. Radwan, M. S. El-Hennawey, and R. A. Goubran, "Non-intrusive single-ended speech quality assessment in VoIP," *Speech Commun.*, vol. 49, no. 6, pp. 477–489, Jun. 2007.
- [100] T. H. Falk and W.-Y. Chan, "Hybrid signal-and-link-parametric speech quality measurement for VoIP communications," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 16, no. 8, pp. 1579–1589, Nov. 2008.
- [101] C. Jin and R. Kubichek, "Output-based objective speech quality using vector quantization techniques," in *ASILOMAR. Conf. on Signals, Systems and Computers*, 1995, vol. 2, pp. 1291–1294.
- [102] W. Li and R. Kubichek, "Output-based objective speech quality measurement using continuous Hidden Markov Models," in *Seventh International Symposium on Signal Processing and Its Applications*, 2003, pp. 389 – 392.
- [103] G. Talwar and R. Kubichek, "Output based speech quality measurement using hidden Markov models," in *International Signal Process. Conference*, 2003.
- [104] L. Roychoudhuri and E. Al-Shaer, "Real-time audio quality evaluation for adaptive multimedia protocols," *Manag. Multimed. Networks Serv.*, vol. 3754, 2005.
- [105] Y. Jung, A. A. Ibanez, and C. Manzano, "Simple QoE estimation model for low capacity VoIP phones," *Electron. Lett.*, vol. 47, no. 3, p. 184, 2011.
- [106] I.-H. Mkwawa, E. Jammeh, and L. Sun, "Mapping of Received Signal Strength Indicator to QoE in VoIP applications over wlan," in *2012 Fourth International Workshop on Quality of Multimedia Experience*, 2012, pp. 156–157.
- [107] J. A. Hassan, M. Hassan, S. K. Das, and A. Ramer, "Managing Quality of Experience for wireless VOIP using noncooperative games," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 7, pp. 1193–1204, Aug. 2012.
- [108] S. Zeadally, H. Moustafa, and F. Siddiqui, "Internet Protocol Television (IPTV): architecture, trends, and challenges," *IEEE Syst. J.*, vol. 5, no. 4, pp. 518–527, Dec. 2011.
- [109] S. Park and S.-H. Jeong, "Mobile IPTV: approaches, challenges, standards, and QoS support," *IEEE Internet Comput.*, vol. 13, no. 3, pp. 23–31, May 2009.

-
- [110] J. Krikke, "Streaming video transforms the media industry," *IEEE Comput. Graph. Appl.*, vol. 24, no. 4, pp. 6–12, Jul. 2004.
 - [111] M. Sadiku and S. Nelatury, "IPTV: an alternative to traditional cable and satellite television," *IEEE Potentials*, vol. 30, no. 4, pp. 44–46, Jul. 2011.
 - [112] Alexandra Suich, "YourTube. The Internet is changing television habits," *The Economist (in the World print edition)*.
 - [113] "Internet protocol television," *The Economist (Economic and Financial Indicators)*.
 - [114] A. Takahashi, D. Hands, and V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," *IEEE Commun. Mag.*, vol. 46, no. 2, pp. 78–84, Feb. 2008.
 - [115] A. Raake, J. Gustafsson, S. Argyropoulos, M. Garcia, D. Lindegren, G. Heikkila, M. Pettersson, P. List, and B. Feiten, "IP-based mobile and fixed network audiovisual media services," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 68–79, Nov. 2011.
 - [116] "Parametric non-intrusive assessment of audiovisual media streaming quality," *ITU-T Recomm. P.1201*, 2012.
 - [117] "Parametric non-intrusive bitstream assessment of video media streaming quality," *ITU-T Recomm. P.1202*, 2012.
 - [118] "Estimating end-to-end performance in IP networks for data applications," *ITU-T Recomm. G.1030*, 2005.
 - [119] "Quality of experience requirements for IPTV services," *ITU-T Recomm. G.1080*, 2008.
 - [120] "Performance monitoring points for IPTV," *ITU-T Recomm. G.1081*, 2008.
 - [121] "Measurement-based methods for improving the robustness of IPTV performance," *ITU-T Recomm. G.1082*, 2009.
 - [122] F. Molazem Tabrizi, J. Peters, and M. Hefeeda, "Dynamic control of receiver buffers in mobile video streaming systems," *IEEE Trans. Mob. Comput.*, vol. 12, no. 5, pp. 995–1008, May 2013.
 - [123] C. Lai, R. Hwang, H. Chao, M. Hassan, and A. Alamri, "A buffer-aware HTTP live streaming approach for SDN-enabled 5G wireless networks," *IEEE Netw.*, vol. 29, no. 1, pp. 49–55, Jan. 2015.
 - [124] M. Sung, "A rich media framework for communication-broadcasting converged IPTV," *IEEE Trans. Consum. Electron.*, vol. 56, no. 2, pp. 463–470, May 2010.

- [125] V. Gopalakrishnan, R. Jana, R. Knag, K. K. Ramakrishnan, D. F. Swayne, and V. A. Vaishampayan, "Characterizing interactive behavior in a large-scale operational IPTV environment," in *IEEE INFOCOM*, 2010, pp. 1–5.
- [126] "Measurements in IP networks for inter-domain performance assessment," *ITU-T Recomm. Y.1543*, 2007.
- [127] "Multicast IP performance parameters," *ITU-T Recomm. Y.1544*, 2008.
- [128] G. Gardikis, L. Boula, G. Xilouris, A. Kourtis, E. Pallis, M. Sidibe, and D. Negru, "Cross-layer monitoring in IPTV networks," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 76–84, Jul. 2012.
- [129] N.-F. Huang, M.-H. Wang, T.-C. Wang, and S.-S. Peng, "Measuring QoE/QoS of large scale P2P IPTV service," in *13th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, 2011, pp. 1–8.
- [130] F. J. Rivas Tocado, A. Diaz Zayas, and P. Merino Gomez, "Characterizing traffic performance in cellular networks," *IEEE Internet Comput.*, vol. 18, no. 1, pp. 12–19, Jan. 2014.
- [131] J. Chen, J. Ma, N. Zhong, Y. Yao, J. Liu, R. Huang, W. Li, Z. Huang, Y. Gao, and J. Cao, "WaaS: Wisdom as a Service," *IEEE Intell. Syst.*, vol. 29, no. 6, pp. 40–47, Nov. 2014.
- [132] "Parametric non-intrusive assessment of audiovisual media streaming quality - lower resolution application area," *ITU-T Recomm. P.1201.1*, 2012.
- [133] "Parametric non-intrusive assessment of audiovisual media streaming quality - higher resolution application area," *ITU-T Recomm. P.1201.2*, 2012.
- [134] "Parametric non-intrusive bitstream assessment of video media streaming quality - Lower resolution application area," *ITU-T Recomm. P.1202.1*, 2012.
- [135] "Parametric non-intrusive bitstream assessment of video media streaming quality - higher resolution application area," *ITU-T Recomm. P.1202.2*, 2013.
- [136] L. Lu, Z. Wang, A. C. Bovik, and J. Kouloheris, "Full-reference video quality assessment considering structural distortion and no-reference quality evaluation of MPEG video," in *IEEE International Conference on Multimedia and Expo*, 2002, vol. 1, pp. 61–64.
- [137] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal Process. Image Commun.*, vol. 19, no. 2, pp. 121–132, 2004.
- [138] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.

-
- [139] S.-O. Lee and D.-G. Sim, "New full-reference visual quality assessment based on human visual perception," in *Digest of Technical Papers - International Conference on Consumer Electronics*, 2008, pp. 1–2.
- [140] "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," *ITU-T Recomm. J.144*, 2004.
- [141] A. Bhat, I. Richardson, and S. Kannangara, "A new perceptual quality metric for compressed video," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 933–936.
- [142] "Reference algorithm for computing peak signal to noise ratio of a processed video sequence with compensation for constant spatial shifts, constant temporal shift, and constant luminance gain and offset," *ITU-T Recomm. J.340*, 2010.
- [143] W. Z. Tobias Oelbaum, Klaus Diepold, "A generic method to increase the prediction accuracy of visual quality metrics," in *Picture Coding Symposium (PCS)*, 2007.
- [144] F. Agboma and A. Liotta, "User centric assessment of mobile contents delivery," in *Fourth International Conference on Advances in Mobile Computing and Multimedia (MoMM)*, 2006, pp. 121–130.
- [145] K. Yamagishi and T. Hayashi, "Parametric packet-layer model for monitoring video quality of IPTV services," in *IEEE International Conference on Communications*, 2008, pp. 110–114.
- [146] J. Joskowicz and J. C. L. Ardao, "Enhancements to the opinion model for video-telephony applications," in *5th International Latin American Networking Conference- LANC '09*, 2009, p. 87.
- [147] F. You, W. Zhang, and J. Xiao, "Packet loss pattern and parametric video quality model for IPTV," in *IEEE/ACIS 8th International Conference on Computer and Information Science*, 2009, pp. 824–828.
- [148] T. Yamada, S. Yachida, Y. Senda, and M. Serizawa, "Accurate video-quality estimation without video decoding," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 2426–2429.
- [149] A. Raake, M.-N. Garcia, S. Moller, J. Berger, F. Kling, P. List, J. Johann, and C. Heidemann, "T-V-model: parameter-based prediction of IPTV quality," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 1149–1152.
- [150] M.-N. Garcia, P. List, S. Argyropoulos, D. Lindegren, M. Pettersson, B. Feiten, J. Gustafsson, and A. Raake, "Parametric model for audiovisual quality assessment in IPTV: ITU-T Rec. P.1201.2," in *IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*, 2013, pp. 482–487.

- [151] K. Yamagishi and G. Shan, "Packet-header-based-quality-estimation model for mobile audiovisual media streaming – Recommendation ITU-T P.1201.1," in *7th International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, 2013, pp. 52–57.
- [152] "Opinion model for video-telephony applications," *ITU-T Recomm. G.1070*, 2012.
- [153] D. De Vera, P. Rodriguez-Bocca, and G. Rubino, "QoE Monitoring Platform for Video Delivery Networks," *IP Oper. Manag.*, vol. 4786, pp. 131–142, 2007.
- [154] S. Winkler and P. Mohandas, "The evolution of video quality measurement: from PSNR to hybrid metrics," *IEEE Trans. Broadcast.*, vol. 54, no. 3, pp. 660–668, Sep. 2008.
- [155] J. Joskowicz and J. Ardao, "Combining the effects of frame rate, bit rate, display size and video content in a parametric video quality model," in *6th Latin America Networking Conference*, 2011, pp. 4–11.
- [156] S. Argyropoulos, A. Raake, M.-N. Garcia, and P. List, "No-reference bit stream model for video quality assessment of h.264/AVC video based on packet loss visibility," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 1169–1172.
- [157] G. Rubino, P. Tirilly, and M. Varela, "Evaluating users' satisfaction in packet networks using random neural networks," in *The International Conference on Artificial Neural Networks (ICANN)*, 2006, pp. 303–312.
- [158] M. Barkowsky, N. Staelens, and L. Janowski, "Open collaboration on hybrid video quality models - VQEG joint effort group hybrid," in *IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*, 2013, pp. 476–481.
- [159] K. Yamagishi, T. Kawano, and T. Hayashi, "Hybrid video-quality-estimation model for IPTV services," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2009, pp. 1–5.
- [160] A. Khan, L. Sun, and E. Ifeachor, "An ANFIS-based hybrid video quality prediction model for video streaming over wireless networks," in *2nd International Conference on Next Generation Mobile Applications, Services, and Technologies*, 2008, pp. 357–362.
- [161] O. Sugimoto, S. Naito, S. Sakazawa, and A. Koike, "Objective perceptual video quality measurement method based on hybrid no reference framework," in *16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 2237–2240.
- [162] M. C. Q. Farias, M. M. Carvalho, H. T. M. Kussaba, and B. H. A. Noronha, "A hybrid metric for digital video quality assessment," in *IEEE International*

- Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2011, pp. 1–6.
- [163] C. Keimel, J. Habigt, and K. Diepold, “Hybrid no-reference video quality metric based on multiway PLSR,” in *20th European Signal Processing Conference (EUSIPCO)*, 2012, pp. 1244–1248.
- [164] H. Minkowski, “Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern,” *Math. Ann.*, vol. 68, no. 4, pp. 472–525, Dec. 1910.
- [165] D. Geerts, K. De Moor, I. Ketyko, A. Jacobs, J. Van den Bergh, W. Joseph, L. Martens, and L. De Marez, “Linking an integrated framework with appropriate methods for measuring QoE,” in *2nd International Workshop on Quality of Multimedia Experience (QoMEX)*, 2010, pp. 158–163.
- [166] S. Moller, K.-P. Engelbrecht, C. Kuhnel, I. Wechsung, and B. Weiss, “A taxonomy of quality of service and Quality of Experience of multimodal human-machine interaction,” in *International Workshop on Quality of Multimedia Experience*, 2009, pp. 7–12.
- [167] K. Kilkki, “Quality of Experience in Communications Ecosystem,” *J. Univers. Comput. Sci.*, vol. 14, no. 5, pp. 615–624, 2008.
- [168] H. H. Song, Z. Ge, A. Mahimkar, J. Wang, J. Yates, Y. Zhang, A. Basso, and M. Chen, “Q-Score: proactive service quality assessment in a large IPTV system,” in *ACM SIGCOMM conference on Internet Measurement Conference (IMC)*, 2011, p. 195.
- [169] K. B. Robert Kooij, Kamal Ahmed, “Perceived Quality of Channel Zapping,” in *Fifth IAESTED International Conference on Communication Systems and Networks (CSN)*, 2006, pp. 155–158.
- [170] B. E. Godana, R. E. Kooij, and O. K. Ahmed, “Impact of advertisements during channel zapping on Quality of Experience,” in *Fifth International Conference on Networking and Services (ICNS)*, 2009, pp. 249–254.
- [171] D. Manzato and N. da Fonseca, “A survey of channel switching schemes for IPTV,” *IEEE Commun. Mag.*, vol. 51, no. 8, pp. 120–127, Aug. 2013.
- [172] F. Ramos, “Mitigating IPTV zapping delay,” *IEEE Commun. Mag.*, vol. 51, no. 8, pp. 128–133, Aug. 2013.
- [173] A. Azgin and Y. Altunbasak, “Dynamic channel reordering to reduce latency during surfing periods in IPTV networks,” *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 471–483, Sep. 2013.
- [174] Eunji Lee, Jung Ku, and Hyokyung Bahn, “An efficient hot channel identification scheme for IPTV channel navigation,” *IEEE Trans. Consum. Electron.*, vol. 60, no. 1, pp. 124–129, Feb. 2014.

- [175] T. Hayashi and A. Takahashi, "QoE assessment method for video quality and pricing in IPTV services," in *ETSI workshop on Effects of Transmission Performance on Multimedia QoS*, 2008.
- [176] P. Reichl, P. Maille, P. Zwickl, and A. Sackl, "A fixed-point model for QoE-based charging," in *Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking - FhMN '13*, 2013, p. 33.
- [177] "Jitsi." [Online]. Available: <http://www.jitsi.org>.
- [178] "Skype." [Online]. Available: <http://www.skype.com>.
- [179] "Wireshark," 2015. [Online]. Available: <http://www.wireshark.org>.
- [180] D. M. Diez, C. D. Barr, and M. Cetinkaya-Rundel, *OpenIntro Statistics*, 2nd ed. 2014.
- [181] "GAWK: effective AWK programming: a user's guide for GNU AWK," 2014. [Online]. Available: <http://www.gnu.org/software/gawk/manual/>.
- [182] "One Way transmission time," *ITU-T Recomm. G.114*, 2003.
- [183] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. Characterization," *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 90–100, Jul. 1997.
- [184] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, no. 24, pp. 46–156, 1945.
- [185] M. Nakagami, "The m-Distribution, a general formula of intensity of rapid fading," in *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium held at the University of California*, 1960, pp. 3–36.
- [186] H. T. Friis, "A note on a simple transmission formula," in *IRE'46*, 1946, vol. 34, no. 5, pp. 254–256.
- [187] G. H. Spencer and M. V. R. K. Murty, "General ray-tracing procedure," *Journal of the Optical Society of America A: Optics, Image Science, and Vision*. Optical Society of America, 01-Jun-1962.
- [188] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.
- [189] Wikipedia, "No Title." [Online]. Available: <http://www.wikipedia.org/>.
- [190] S. Shin and H. Schulzrinne, "Measurement and analysis of the VoIP capacity in IEEE 802.11 WLAN," *IEEE Trans. Mob. Comput.*, vol. 8, no. 9, pp. 1265–1279, Sep. 2009.

- [191] K. Medepalli, P. Gopalakrishnan, D. Famolari, and T. Kodama, "Voice capacity of IEEE 802.11b, 802.11a and 802.11g wireless LANs," in *GLOBECOM '04*, 2004, vol. 3, pp. 1549–1553.
- [192] D. P. Hole and F. A. Tobagi, "Capacity of an IEEE 802.11b wireless LAN supporting VoIP," in *ICC'04*, 2004, pp. 196–201.
- [193] S. Garg and M. Kappes, "An experimental study of throughput for UDP and VoIP traffic in IEEE 802.11b networks," in *WCNC'03*, 2003, vol. 3, pp. 1748–1753.
- [194] A. Lakas and M. Boulmalf, "Experimental analysis of VoIP over wireless local area networks," *J. Commun.*, vol. 2, no. 4, pp. 3–9, Jun. 2007.
- [195] L. Cai, X. Shen, J. W. Mark, and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, no. 3, pp. 752–761, May 2006.
- [196] S. Harsha, A. Kumar, and V. Sharma, "An analytical model for the capacity estimation of combined VoIP and TCP file transfers over EDCA in an IEEE 802.11e WLAN," in *IWQoS'06*, 2006, pp. 178–187.
- [197] "Pulse Code Modulation (PCM) of voice frequencies," *ITU-T Recomm. G.711*, 1972.
- [198] O. Awoniyi and F. A. Tobagi, "Effect of fading on the performance of VOIP in IEEE 802.11a WLANs," in *ICC'04*, 2004, pp. 3712–3717.
- [199] S. Aksu, E. E. Gungor, and G. Karabulut Kurt, "Effect of Nakagami-m fading on the QoE performance of VoIP in wireless mesh networks," in *WoWMoM'11*, 2011, pp. 1–6.
- [200] A. H. Reeves, "Electric signaling system," 2272070, 1942.
- [201] D. W. Petr, "Adaptive differential PCM coding," 4437087, 1984.
- [202] Y. Liu, H. Du, Y. Q. J. Xu, Y. Ye, X. Zhou, Y. Wang, K. Xu, X. Meng, W. Hu, and X. Li, "Method for evaluating QoE (Quality Of Experience) of voice service user perception experience by simulating user behaviors," 102685789, 2012.
- [203] J. Cao, H. Qian, and J. Li, "Quality-of-experience measurement for voice services," 20130051271, 2011.
- [204] A. Neumann, C. Aichele, M. Lindner, and S. Wunderlich, "Better Approach To Mobile Ad-hoc Networking (B.A.T.M.A.N.)," *IETF Draft*, 2008.
- [205] J. Pojda, A. Wolff, M. Sbeiti, and C. Wietfeld, "Performance analysis of mesh routing protocols for UAV swarming applications," *8th Int. Symp. Wirel. Commun. Syst.*, pp. 317–321, Nov. 2011.

- [206] M. Reineri, C. Casetti, and C.-F. Chiasserini, "Routing protocols for mesh networks with mobility support," in *6th International Symposium on Wireless Communication Systems*, 2009, pp. 71–75.
- [207] T. Clausen and P. Jacquet, "Optimized Link State Routing (OLSR) Protocol," *IETF RFC 3626*, 2003.
- [208] R. K. Gujral and M. Singh, "Performance analysis of ad hoc routing protocols for voice communication support over hybrid MANETs," *Int. J. Comput. Appl.*, vol. 22, no. 3, pp. 42–48, 2011.
- [209] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing," *IETF RFC 3561*, 2003.
- [210] K. L. Espensen, M. K. Kjeldsen, and L.M.Kristensen, "Modelling and initial validation of the DYMO routing protocol for mobile ad-hoc networks," *Lect. Notes Comput. Sci.*, vol. 5062, pp. 152–170, 2008.
- [211] V. Park and S. Corson, "Temporally-Ordered Routing Algorithm (TORA)," *IETF Draft*, 1999.
- [212] Z. J. Haas, M. R. Pearlman, and P. Samar, "The Zone Routing Protocol (ZRP) for ad hoc networks," *IETF Draft*, 2002.
- [213] G. Adam, C. Bouras, A. Gkamas, V. Kapoulas, G. Kioumourtzis, and N. Tavoularis, "Performance evaluation of routing protocols for multimedia transmission over mobile ad hoc networks," in *4th Joint IFIP Wireless and Mobile Networking Conference (WMNC 2011)*, 2011, pp. 1–6.
- [214] P. Manjul, V. Balasubramanian, Y. Li, Y. Shi, Y. Liu, J. Xu, Q. Xie, J. Deng, H. Li, and C.-K. Siew, "Real-time video streaming over multi-hop ad-hoc networks," in *Second International Conference on Networking and Distributed Computing*, 2011, pp. 283–288.
- [215] D. Johnson, Y. Hu, and D. Maltz, "The Dynamic Source Routing Protocol (DSR) for mobile ad hoc networks for IPv4," *IETF RFC 4728*, 2007.
- [216] R. Vaishampayan and J. J. Garcia-Luna-Aceves, "Efficient and robust multicast routing in mobile ad hoc networks," in *IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, 2004, pp. 304–313.
- [217] F. Community, "Better Approach To Mobile Ad hoc Networking (B.A.T.M.A.N.)." [Online]. Available: <<http://www.open-mesh.org>>.
- [218] D. Johnson, N. Ntlatlapa, and C. Aichele, "A simple pragmatic approach to mesh routing using BATMAN," in *2nd IFIP International Symposium on Wireless Communications and Information Technology in Developing Countries*, 2008.
- [219] E. Kulla, M. Ikeda, T. Oda, L. Barolli, F. Xhafa, and M. Takizawa, "Multimedia Transmissions over a MANET Testbed: Problems and Issues," in *Sixth*

- International Conference on Complex, Intelligent, and Software Intensive Systems*, 2012, pp. 141–147.
- [220] D. Seither, A. König, and M. Hollick, “Routing performance of wireless mesh networks: a practical evaluation of BATMAN advanced,” in *IEEE 36th Conference on Local Computer Networks*, 2011, pp. 897–904.
 - [221] C. Siaterlis, A. P. Garcia, and B. Genge, “On the use of Emulab testbeds for scientifically rigorous experiments,” *IEEE Commun. Surv. Tutorials*, vol. 15, no. 2, pp. 929–942, Jan. 2013.
 - [222] V. Asghari, D. B. da Costa, and S. Aissa, “Performance analysis for multihop relaying channels with Nakagami-m fading: ergodic capacity upper-bounds and outage probability,” *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2761–2767, Oct. 2012.
 - [223] D. Torrieri and M. C. Valenti, “The outage probability of a finite ad Hoc network in Nakagami fading,” *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3509–3518, Nov. 2012.
 - [224] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. J. Paulraj, “Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, Jun. 2010.
 - [225] A. Botta, A. Dainotti, and A. Pescapé, “A tool for the generation of realistic network workload for emerging networking scenarios,” *Comput. Networks*, vol. 56, no. 15, pp. 3531–3547, Oct. 2012.
 - [226] L. Angrisani, D. Capriglione, L. Ferrigno, and G. Miele, “An Internet Protocol packet delay variation estimator for reliable quality assessment of video-streaming services,” *IEEE Trans. Instrum. Meas.*, vol. 62, no. 5, pp. 914–923, May 2013.
 - [227] B. Barekatain, M. A. Maarof, A. A. Quintana, and A. T. Cabrera, “GREENIE: a novel hybrid routing protocol for efficient video streaming over wireless mesh networks,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2013, no. 1, p. 168, Jun. 2013.
 - [228] I. G. Tudorache, A. M. Popescu, and A. H. Kemp, “MANET routing protocols problem for the marginal mobility model,” in *41st European Microwave Conference (EuMA)*, 2011, pp. 139–142.
 - [229] D. Helbing, “Traffic and related self-driven many-particle systems,” *Rev. Mod. Phys.*, vol. 73, no. 4, pp. 1067–1141, Dec. 2001.
 - [230] A. Nascimento, S. Queiroz, L. Galvao, E. Mota, and E. Nascimento, “Influence of propagation modeling on VoIP quality performance in wireless mesh network simulation,” in *IEEE International Symposium on Modeling, Analysis and Simulation of Computers and Telecommunication Systems*, 2008, pp. 1–3.

- [231] P. K. Singh, "Influences of TwoRayGround and Nakagami propagation model for the performance of adhoc routing protocol in VANET," *Int. J. Comput. Appl.*, vol. 45, no. 22, pp. 1–6, 2012.
- [232] E. Modiano, "An adaptive algorithm for optimizing the packet size used in wireless ARQ protocols," *Wirel. Networks*, vol. 5, pp. 279–286, 1999.
- [233] D. Murray, M. Dixon, and T. Koziniec, "An experimental comparison of routing protocols in multi hop ad hoc networks," in *Australasian Telecommunication Networks and Applications Conference*, 2010, pp. 159–164.
- [234] "University of Utah Flux Research Group, Emulab: The Utah Network Testbed." [Online]. Available: <http://www.emulab.net>.
- [235] Y. Huang, S. Bhatti, and D. Parker, "Tuning OLSR," in *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2006, pp. 1–5.
- [236] S. Demers and L. Kant, "MANETs: performance analysis and management," in *MILCOM '06*, 2006, pp. 1–7.
- [237] M. Hiyama, E. Kulla, M. Ikeda, L. Barolli, and M. Takizawa, "Investigation of OLSR behavior for different Hello packets intervals in a MANET testbed," in *IEEE 27th International Conference on Advanced Information Networking and Applications (AINA)*, 2013, pp. 183–188.
- [238] S. Liu and A. D. Striegel, "Exploring the potential in practice for opportunistic networks amongst smart mobile devices," in *19th annual International Conference on Mobile computing & networking - MobiCom '13*, 2013, p. 315.
- [239] A. Boukerche and A. Darehshoorzadeh, "Opportunistic routing in wireless networks: models, algorithms, and classifications," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 1–36, Nov. 2014.
- [240] G. P. Perrucci, F. H. P. Fitzek, and J. Widmer, "Survey on energy consumption entities on the smartphone platform," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1–6.
- [241] S. Chiaravalloti, F. Idzikowski, and Ł. Budzisz, "Power consumption of WLAN network elements," *TKN Technical Report TKN-11-002. TU Berlin*. 2011.
- [242] "Information technology - MPEG Systems Technologies - Part 11: Energy-Efficient Media Consumption (Green Metadata), ISO/IEC JTC1/SC29/WG11 Final Draft International Standard 23001-11." 2014.
- [243] F. C. Fernandes, X. Ducloux, Z. Ma, E. Faramarzi, P. Gendron, and J. Wen, "The Green Metadata Standard for energy-efficient video consumption," *IEEE Multimed.*, vol. 22, no. 1, pp. 80–87, Jan. 2015.

-
- [244] C. Singhal, S. De, R. Trestian, and G.-M. Muntean, "Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast," *IEEE Trans. Mob. Comput.*, vol. 13, no. 7, pp. 1522–1535, Jul. 2014.
- [245] C. Singhal, S. De, R. Trestian, and G.-M. Muntean, "eWU-TV: user-centric energy-efficient digital TV broadcast over Wi-Fi networks," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 39–55, Mar. 2015.
- [246] S. Wang, W. Guo, C. Khirallah, D. Vukobratovic, and J. Thompson, "Interference allocation scheduler for green multimedia delivery," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2059–2070, Jun. 2014.
- [247] Y. Xu, R. Q. Hu, L. Wei, and G. Wu, "QoE-aware mobile association and resource allocation over wireless heterogeneous networks," in *IEEE Global Communications Conference*, 2014, pp. 4695–4701.
- [248] Z. Tang, Z. Wang, P. Li, S. Guo, X. Liao, and H. Jin, "An application layer protocol for energy-efficient bandwidth aggregation with Guaranteed Quality-of-Experience," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 6, pp. 1538–1546, Jun. 2015.
- [249] E. Danish, V. Silva, A. Fernando, C. Alwis, and A. Kondo, "Content-aware resource allocation in OFDM systems for energy-efficient video transmission," *IEEE Trans. Consum. Electron.*, vol. 60, no. 3, pp. 320–328, Aug. 2014.
- [250] S. Khan, D. Schroeder, A. El Essaili, and E. Steinbach, "Energy-efficient and QoE-driven adaptive HTTP streaming over LTE," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 2354–2359.
- [251] Shivashankar, H. N. Suresh, G. Varaprasad, and G. Jayanthi, "Designing Energy Routing Protocol With Power Consumption Optimization in MANET," *IEEE Trans. Emerg. Top. Comput.*, vol. 2, no. 2, pp. 192–197, Jun. 2014.
- [252] S. Lim, C. Yu, and C. R. Das, "RandomCast: an energy-efficient communication scheme for mobile ad hoc networks," *IEEE Trans. Mob. Comput.*, vol. 8, no. 8, pp. 1039–1051, Aug. 2009.
- [253] F. De Rango, F. Guerriero, and P. Fazio, "Link-stability and energy aware routing protocol in distributed wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 4, pp. 713–726, Apr. 2012.
- [254] L. Chen and W. B. Heinzelman, "QoS-aware routing based on bandwidth estimation for mobile ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 3, pp. 561–572, Mar. 2005.
- [255] T. Lu and J. Zhu, "Genetic algorithm for energy-efficient QoS multicast routing," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 31–34, Jan. 2013.

- [256] W. L. Huang and K. Ben Letaief, "Cross-layer scheduling and power control combined with adaptive modulation for wireless ad hoc networks," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 728–739, Apr. 2007.
- [257] B. Tavli and W. B. Heinzelman, "Energy and spatial reuse efficient network-wide real-time data broadcasting in mobile ad hoc networks," *IEEE Trans. Mob. Comput.*, vol. 5, no. 10, pp. 1297–1312, Oct. 2006.
- [258] B. Tavli and W. B. Heinzelman, "Energy-efficient real-time multicast routing in mobile ad hoc networks," *IEEE Trans. Comput.*, vol. 60, no. 5, pp. 707–722, May 2011.
- [259] J.-L. Kuo, C.-H. Shih, C.-Y. Ho, and Y.-C. Chen, "A cross-layer approach for real-time multimedia streaming on wireless peer-to-peer ad hoc network," *Ad Hoc Networks*, vol. 11, no. 1, pp. 339–354, Jan. 2013.
- [260] Y. Peng, Y. Yu, L. Guo, D. Jiang, and Q. Gai, "An efficient joint channel assignment and QoS routing protocol for IEEE 802.11 multi-radio multi-channel wireless mesh networks," *J. Netw. Comput. Appl.*, vol. 36, no. 2, pp. 843–857, Mar. 2013.
- [261] C. Westphal, "Opportunistic routing in dynamic ad hoc networks: the OPRAH protocol," in *IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, 2006, pp. 570–573.
- [262] M. Zorzi and R. R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Trans. Mob. Comput.*, vol. 2, no. 4, pp. 337–348, Oct. 2003.
- [263] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *9th annual international conference on Mobile computing and networking - MobiCom '03*, 2003, p. 134.
- [264] A. Darehshoorzadeh and L. Cerda-Alabern, "Distance progress based opportunistic routing for wireless mesh networks," in *8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2012, pp. 179–184.
- [265] Z. Zhong, J. Wang, S. Nelakuditi, and G.-H. Lu, "On selection of candidates for opportunistic anypath forwarding," *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 10, no. 4, p. 1, Oct. 2006.
- [266] P. Larsson, "Selection diversity forwarding in a multihop packet radio network with fading channel and capture," *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 5, no. 4, p. 47, Oct. 2001.
- [267] S. Biswas and R. Morris, "ExOR: opportunistic multi-hop routing for wireless networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 4, p. 133, Oct. 2005.

-
- [268] R. Ahlswede, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [269] S.-C. Lin and K.-C. Chen, "Spectrum-Map-Empowered opportunistic routing for cognitive radio ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2848–2861, Jul. 2014.
- [270] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 4, p. 169, Oct. 2007.
- [271] P. Spachos and D. Hantzinakos, "Scalable dynamic routing protocol for cognitive radio sensor networks," *IEEE Sens. J.*, vol. 14, no. 7, pp. 2257–2266, Jul. 2014.
- [272] P. Li, S. Guo, S. Yu, and A. V. Vasilakos, "CodePipe: an opportunistic feeding and routing protocol for reliable multicast with pipelined network coding," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3264–3273, 2014.
- [273] H. Shen, G. Bai, Z. Tang, and L. Zhao, "Flow-correlation-aware opportunistic network coding for video streaming over wireless networks," in *8th International Conference on Communications and Networking in China (CHINACOM)*, 2013, pp. 155–160.
- [274] K. Choumas, I. Syrigos, T. Korakis, and L. Tassiulas, "Video-aware multicast opportunistic routing over 802.11 two-hop mesh networks," in *Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2014, pp. 486–494.
- [275] D. Koutsonikolas, Y. C. Hu, and C.-C. Wang, "Pacifier: high-throughput, reliable multicast without 'crying babies' in wireless mesh networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 5, pp. 1375–1388, Oct. 2012.
- [276] H. Seferoglu and A. Markopoulou, "Opportunistic network coding for video streaming over wireless," in *Packet Video 2007*, 2007, pp. 191–200.
- [277] F. Wu, K. Gong, T. Zhang, G. Chen, and C. Qiao, "COMO: a game-theoretic approach for joint multirate opportunistic routing and forwarding in non-cooperative wireless networks," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 2, pp. 948–959, Feb. 2015.
- [278] S. Wang, A. Basalamah, S. M. Kim, S. Guo, Y. Tobe, and T. He, "Link-correlation-aware opportunistic routing in wireless networks," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 1, pp. 47–56, Jan. 2015.
- [279] J. Luo, J. Hu, D. Wu, and R. Li, "Opportunistic routing algorithm for relay node selection in wireless sensor networks," *IEEE Trans. Ind. Informatics*, vol. 11, no. 1, pp. 112–121, Feb. 2015.

- [280] L. Cheng, J. Niu, J. Cao, S. K. Das, and Y. Gu, "QoS aware geographic opportunistic routing in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 7, pp. 1864–1875, Jul. 2014.
- [281] TP-Link, "Link budget calculation." [Online]. Available: <http://www.tp-link.com/en/support/calculator/>.
- [282] UltraElectronics and Wood&Douglas, "What is fade margin." [Online]. Available: <http://ultra-woodanddouglas.com/FAQ/what-is-fade-margin>.
- [283] "Video Trace Library." [Online]. Available: <http://trace.eas.asu.edu>.
- [284] Broadcom, "BCM4330 Preliminary Data Sheet," 2011.
- [285] L. M. Feeney and D. Willkomm, "Energy Framework: an extensible framework for simulating battery consumption in wireless networks," in *3rd International ICST Conference on Simulation Tools and Techniques*, 2012, p. 20.
- [286] F. B. Abdesslem, L. Iannone, M. Dias de Amorim, K. Kabassanov, and S. Fdida, "On the feasibility of power control in current IEEE 802.11 devices," in *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06)*, 2006, pp. 468–473.
- [287] M. Bouhorma, A. Boudhir, M. Ben Ahmed, and S. El Brak, "New route request mechanism for energy optimization in mobile ad hoc networks," in *19th Telecommunications Forum (TELFOR)*, 2011, pp. 230–233.
- [288] J. Garcia-Nieto and E. Alba, "Automatic parameter tuning with metaheuristics of the AODV routing protocol for vehicular ad-hoc networks," *Appl. Evol. Comput.*, vol. 6025, pp. 21–30, 2010.
- [289] E. Larsen, J. Flathagen, V. Pham, and L. LandMark, "iOLSR : OLSR for WSNs using dynamically adaptive intervals," in *Fifth International Conference on Sensor Technologies and Applications (SENSORCOMM 2011)*, 2011, pp. 18–23.
- [290] A. Darehshoorzadeh and A. Boukerche, "Opportunistic routing protocols in wireless networks: a performance comparison," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, 2014, pp. 2504–2509.
- [291] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019. White Paper," *White Pap.*, 2015.



Universidad
Politécnica
de Cartagena



etsit
escuela técnica superior de
ingeniería de telecomunicación